

VISIT...

LANZAROTE
Caliente.COM

97УДК 575
ББК 28.04
П 20

Ответственный редактор
доктор химических наук Ю.А. Берлин

Рецензенты:
кандидат химических наук В.Г. Коробко
доктор биологических наук В.А. Гвоздев

Патрушев Л.И.

Экспрессия генов. – М.: Наука, 2000. – 000 с., ил.
ISBN 5-02-001890-2

В монографии рассмотрены современные представления о строении и механизмах функционирования генов прокариот и эукариот, а также основные методы их исследования. Книга состоит из двух частей. В первой части обсуждаются структура генома прокариотических и эукариотических организмов, а также механизмы транскрипции, трансляции, репликации, репарации и их регуляции. Во второй части монографии рассмотрены принципы основных методов, используемых в исследованиях генов. Главное внимание уделено современным методам генной инженерии. Обсуждаются наиболее важные аспекты развития современной молекулярной генетики в исследованиях направленного мутагенеза и белковой инженерии, антисмысловых РНК, рибозимов и дезоксирибозимов, трансгеноза и генотерапии, а также достижения в ДНК-диагностике и ДНК-типировании и изучении генома человека.

Для научных работников, аспирантов и студентов, специализирующихся в области химии и биологии.

ТП-97-П-№ 123

Patrushev L.I.

Gene Expression. – Moscow: Nauka, 2000. – 000 p., il.
ISBN 5-02-001890-2

This monograph consisting of two parts presents a comprehensive view of the structure and functioning mechanisms of genes for prokaryotes and eukaryotes as well as the basic methods used for their studies. Its first part discusses the genome structure of prokaryotic and eukaryotic organisms, also covering the mechanisms of transcription, translation, replication, DNA repair and their regulation. The second part of the monograph deals with the principles of the techniques employed in the gene research. The book devotes particular attention to current genetic engineering tools. It embraces the key aspects of recent developments of molecular genetics studies on site-directed mutagenesis, protein engineering, antisense RNA, ribozymes and deoxyribozymes, transgenesis and gene therapy, DNA diagnostics and genotyping human genome.

The book will act as authoritative references for researchers, postgraduates, and students specializing in chemistry and biology.

ISBN 5-02-001890-2

© Издательство "Наука", 2000

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ РЕДАКТОРА.....	9
ПРЕДИСЛОВИЕ АВТОРА	12
ЧАСТЬ I. МЕХАНИЗМЫ ХРАНЕНИЯ И РЕАЛИЗАЦИИ ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ.....	17
ВВЕДЕНИЕ	18
ГЛАВА 1. ГЕНОМ.....	25
1.1. Гены и хромосомы	28
1.2. Геном прокариот	30
1.2.1. Геном вирусов.....	31
1.2.2. Нуклеоид бактериальной клетки	32
1.2.3. Геном архебактерий	35
1.2.4. Минимальный размер генома одноклеточных организмов	37
1.3. Геном эукариот	40
1.3.1. Последовательности нуклеотидов эукариотического генома	41
1.3.2. Хроматин	45
1.3.3. Роль ДНК-топоизомераз в обеспечении структуры и функционирования хроматина	55
ГЛАВА 2. РЕАЛИЗАЦИЯ ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ ПРИ ЭКСПРЕССИИ ГЕНОВ	60
2.1. Транскрипция	61
2.1.1. ДНК-зависимые РНК-полимеразы	62
2.1.2. Единицы транскрипции (транскриптоны)	72
2.1.3. Этапы транскрипции	80
2.1.4. Хроматин во время транскрипции	112
2.2. Котранскрипционные и посттранскрипционные модификации РНК.....	124
2.2.1. Процессинг РНК у бактерий	125
2.2.2. Редактирование пре-мРНК	131
2.2.3. Другие модификации эукариотических мРНК.....	143
2.2.4. Кэп-связывающий комплекс в роли фактора, сопрягающего основные реакции метаболизма транскриптов РНК- полимеразы II	163
2.3. Функциональная компартментализация ядра	168
2.3.1. Интерфазные хромосомы в ядре	169
2.3.2. Ядрышко	171
2.3.3. Пространственная организация синтеза мРНК	173
2.3.4. Ядерные тельца и домены.....	176
2.3.5. Компартментализованное ядро	178
2.4. Биосинтез белка рибосомами бактерий.....	180
2.4.1. Рибосомы	180

2.4.2.	Этапы биосинтеза белка	187
2.4.3.	Антибиотики, действующие на уровне трансляции	197
2.5.	Трансляция у эукариот	200
2.5.1.	Особенности первичной структуры эукариотических мРНК	201
2.5.2.	Инициация биосинтеза белка эукариотическими рибосомами	202
2.5.3.	Элонгация полипептидных цепей	212
2.5.4.	Терминация трансляции	213
2.5.5.	Трансляция в митохондриях	215
2.5.6.	Трансляция в хлоропластах	219
ГЛАВА 3.	ОСНОВНЫЕ ПУТИ РЕГУЛЯЦИИ ЭКСПРЕССИИ ГЕНОВ	223
3.1.	Регуляция экспрессии генов на уровне транскрипции у прокариот	225
3.1.1.	Регуляция на уровне инициации транскрипции	225
3.1.2.	Регуляция синтеза РНК на уровне элонгации и терминации	229
3.2.	Регуляция экспрессии генов на уровне транскрипции у эукариот	238
3.2.1.	Передача сигнала и вторичные мессенджеры	241
3.2.2.	Механизмы позитивной регуляции транскрипции	259
3.2.3.	Механизмы негативной регуляции транскрипции	297
3.2.4.	Структура хроматина как специфический регулятор экспрессии генов	304
3.2.5.	Импринтинг	318
3.2.6.	Метилирование ДНК в регуляции транскрипции	319
3.3.	Посттранскрипционная регуляция экспрессии генов	330
3.3.1.	Направленный транспорт, внутриклеточная локализация и депонирование мРНК	330
3.3.2.	Сплайсинг РНК в регуляции экспрессии генов	343
3.3.3.	Избирательная деградация мРНК	353
3.4.	Регуляция экспрессии генов на уровне трансляции	358
3.4.1.	Регуляция инициации трансляции	358
3.4.2.	Регуляция элонгации синтеза полипептидных цепей	367
3.4.3.	Регуляция терминации трансляции	370
3.5.	Синтез белков, содержащих остатки селеноцистеина	371
3.6.	Посттрансляционная регуляция экспрессии генов	373
3.6.1.	Последствия фолдинга вновь синтезированных полипептидных цепей	373
3.6.2.	Специфические протеиназы в посттрансляционном процессинге белков	376
3.6.3.	Убиквитин-зависимая система протеолиза в регулируемой деградации белков	377
3.6.4.	Сплайсинг белков	381
3.6.5.	Другие посттрансляционные модификации белков	386
ГЛАВА 4.	ВОСПРОИЗВЕДЕНИЕ ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ	390
4.1.	Репликация ДНК	390

4.1.1.	Белки, участвующие в репликации ДНК	391
4.1.2.	Репликативная вилка <i>E. coli</i> и бактериофага Т4	394
4.1.3.	Особенности функционирования репликативной вилки эукариот	400
4.2.	Регуляция репликации ДНК	403
4.2.1.	Инициация репликации ДНК у <i>E. coli</i> и ее регуляция	405
4.2.2.	Регуляция репликации плазмиды ColE1	414
4.3.	Особенности репликации линейных геномов	421
4.3.1.	Линейные хромосомы бактерий	422
4.3.2.	Репликаторы эукариот	425
4.3.3.	Репликация теломерных участков эукариотических хромосом ...	427
4.3.4.	Пространственная организация синтеза ДНК у эукариот	429
ГЛАВА 5.	ЗАЩИТА ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ	432
5.1.	Мутации	433
5.1.1.	Основные источники мутаций и методы определения мутагенной активности	435
5.1.2.	SOS-мутагенез у бактерий	449
5.1.3.	Мутаторный фенотип	453
5.1.4.	Экспансия ДНК	456
5.1.5.	Адаптивные мутации	460
5.1.6.	Механизмы защиты генома от мутаций	464
5.2.	Репарация ДНК	465
5.2.1.	Основные механизмы репарации поврежденной ДНК	466
5.2.2.	Эксцизионная репарация в клетках животных	468
5.2.3.	Гомологичная рекомбинация в репарации ДНК	481
5.2.4.	Репарация ошибочно спаренных нуклеотидов	483
5.2.5.	Полимераза поли(ADP-рибозы) в репарации ДНК у эукариот	486
5.3.	Альтруистичная ДНК	490
5.3.1.	Парадокс возможности существования многоклеточных организмов	491
5.3.2.	Повышение информационной стабильности генома избыточными последовательностями	495
5.3.3.	Селективная защита генов от мутаций	500
5.3.4.	Высокоупорядоченное расположение летальных генов на хромосомах	508
5.3.5.	Возможный смысл парадокса С	516
ГЛАВА 6.	СОВРЕМЕННАЯ КОНЦЕПЦИЯ ГЕНА	522
ЧАСТЬ II. ИСКУССТВЕННЫЕ ГЕНЕТИЧЕСКИЕ СИСТЕМЫ		531
ГЛАВА 7.	ПРИНЦИПЫ ГЕННОЙ ИНЖЕНЕРИИ	532
7.1.	Основные ферменты, используемые в генной инженерии ...	539
7.1.1.	Рестриктазы и ДНК-метилазы	539
7.1.2.	ДНК- и РНК-лигазы	548
7.1.3.	Ферменты матричного синтеза ДНК и РНК	549
7.1.4.	Другие ферменты	556

7.2.	Векторы.....	558
7.2.1.	Плазмидные векторы	559
7.2.2.	Векторы на основе фага λ	563
7.2.3.	Космиды и фазмиды.....	568
7.2.4.	Сверхъемкие векторы YAC, BAC и PAC	570
7.2.5.	Интегрирующие и челночные (бинарные) векторы.....	574
7.2.6.	Конструирование экспрессирующих векторов и их функционирование	577
7.2.7.	Векторы для переноса ДНК в клетки животных и растений	593
7.3.	Клонотеки генов.....	596
7.3.1.	Получение клонотек генов	596
7.3.2.	Введение рекомбинантных ДНК в клетки	599
7.3.3.	Методы скрининга клонотек генов.....	600
7.4.	Эукариотические системы экспрессии рекомбинантных генов, основанные на культурах клеток.....	602
7.4.1.	Клетки яичников китайских хомячков (линия CHO).....	604
7.4.2.	Клетки мышиной миеломы (линия Sp2/0).....	607
7.4.3.	Клетки селезенки мышей (линия MEL)	608
7.4.4.	Клетки африканской зеленой мартышки (линия COS)	610
7.4.5.	Клетки насекомых, зараженные бакуловирусами	611
7.4.6.	Сравнение эффективности рассмотренных систем экспрессии..	612
7.5.	Бесклеточные белоксинтезирующие системы.....	613
7.5.1.	Прокариотические системы	615
7.5.2.	Эукариотические системы	618
7.5.3.	Проточные системы.....	619
7.6.	Другие современные методы исследования генов	621
7.6.1.	Рестрикционное картирование генов	621
7.6.2.	"Прогулки и прыжки по хромосомам"	622
7.6.3.	S1-картирование РНК и ДНК.....	624
7.6.4.	Футпринтинг	625
7.7.	Стратегия выделения нового гена	626
ГЛАВА 8.	НАПРАВЛЕННЫЙ МУТАГЕНЕЗ И БЕЛКОВАЯ ИНЖЕНЕРИЯ ...	629
8.1.	Методы направленного получения мутаций.....	630
8.1.1.	Получение делеций и вставок	630
8.1.2.	Химический мутагенез.....	633
8.1.3.	Сайт-специфический мутагенез с использованием олигонуклеотидов	635
8.1.4.	Полимеразная цепная реакция в направленном мутагенезе.....	642
8.2.	Белковая инженерия.....	646
8.2.1.	Библиотеки пептидов и эпитопов	646
8.2.2.	Белки-репортеры в гибридных белках	655
8.2.3.	Гибридные токсины	656
8.2.4.	Подходы к созданию новых ферментов.....	662
8.2.5.	Субтилигаза в лигировании пептидов.....	666
8.3.	Концепция ксенобиоза.....	670

ГЛАВА 9. АНТИСМЫСЛОВЫЕ РНК, РИБОЗИМЫ И ДЕЗОКСИРИБОЗИМЫ	679
9.1. Антисмысловые РНК и олигонуклеотиды	679
9.1.1. Механизм действия антисмысловых РНК	680
9.1.2. Использование антисмысловых РНК	682
9.1.3. Природные антисмысловые РНК	687
9.1.4. Антисмысловые РНК и патология: возможный механизм возникновения доминантных мутаций	690
9.2. Рибозимы и дезоксирибозимы	692
9.2.1. Типы рибозимов	692
9.2.2. Свойства рибозимов	694
9.2.3. Рибозимы как лекарственные средства	698
9.2.4. Репарация мутантных РНК с помощью рибозимов, осуществляющих <i>транс</i> -сплайсинг	704
9.2.5. Дезоксирибозимы	708
9.3. Аптамеры	710
9.4. Молекулы РНК у истоков жизни	713
9.4.1. Молекулы РНК в качестве РНК-репликаз	714
9.4.2. Возможность синтеза полипептидных цепей молекулами РНК	717
ГЛАВА 10. ТРАНСГЕННЫЕ ЖИВОТНЫЕ И РАСТЕНИЯ	722
10.1. Способы получения трансгенных многоклеточных организмов	722
10.2. Экспрессия трансгенов	725
10.3. Использование трансгенов у животных	726
10.3.1. Исследование механизмов экспрессии генов	727
10.3.2. Токсигены в исследовании дифференцировки соматических клеток в онтогенезе	728
10.3.3. Изменение физиологического статуса лабораторных и сельскохозяйственных животных	729
10.3.4. Моделирование наследственных и приобретенных заболеваний человека	732
10.4. Трансгенные растения	735
10.5. Генотерапия наследственных и приобретенных заболеваний	737
10.5.1. Способы доставки новых генов в геном человека	738
10.5.2. Управление экспрессией трансгенов в клетках-мишенях	743
10.5.3. Современные достижения генотерапии онкологических заболеваний	746
10.5.4. Ближайшие перспективы использования генотерапии	750
10.5.5. Успехи генотерапии в модельных экспериментах	752
10.5.6. Проблемы, возникающие в связи с практическим применением генотерапии	753
ГЛАВА 11. ДНК-ДИАГНОСТИКА И ДНК-ТИПИРОВАНИЕ	755

11.1.	ДНК-диагностика наследственных и приобретенных заболеваний.....	759
11.1.1.	Получение клинического генетического материала	759
11.1.2.	Диагностика заболеваний.....	761
11.2.	ДНК-типирование.....	774
11.2.1.	ДНК-типирование микроорганизмов	775
11.2.2.	Идентификация личности на основе минисателлитной ДНК: определение отцовства.....	778
11.3.	Микроматрицы и микрочипы ДНК	786
11.3.1.	Методы создания микроматриц ДНК	787
11.3.2.	Ограничения в использовании микроматриц ДНК	790
11.3.3.	Использование микроматриц ДНК в фундаментальных и прикладных исследованиях	792
ГЛАВА 12.	КАРТИРОВАНИЕ И ОПРЕДЕЛЕНИЕ ПЕРВИЧНОЙ СТРУКТУРЫ ГЕНОМА ЧЕЛОВЕКА	796
12.1.	Основные подходы к картированию генома человека	796
12.1.1.	Генетические карты сцепления.....	797
12.1.2.	ПЦР в исследованиях генома человека	802
12.1.3.	Физические карты низкого разрешения	804
12.1.4.	Физические карты высокого разрешения	807
12.2.	Определение полной первичной структуры ДНК генома человека	809
12.3.	Базы данных получаемой информации.....	810
	ЗАКЛЮЧЕНИЕ	813
	РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА	818

Предисловие редактора

*Сей факт с сияющим лицом
Вношу как ценный вклад в науку.
Саша Черный*

Науку двигают методы и увенчивают теории. Концептуальные прорывы 1950–1960-х годов – прежде всего, двойная спираль, генетический код и механизм биосинтеза белка – легли в основу молекулярной биологии и молекулярной генетики. Этот начальный период бури и натиска завершился химико-ферментативным синтезом структурного гена аланиновой тРНК дрожжей – событием, которое далеко не все тогда оценили по достоинству. Для скудных методических возможностей того времени это был феноменальный *coup de force*, потребовавший пятилетних усилий группы высококвалифицированных химиков и биохимиков под руководством Гобинда Кораны, чтобы быть выполненным, и целого номера "Journal of Molecular Biology", чтобы быть описанным. Шокированный последним обстоятельством, журнал "Nature New Biology" устами своего рецензента заявил: «Подобно проекту "Аполлон", все это сделано лишь для того чтобы показать, что это можно сделать, и, как и этот проект, никогда не будет повторено». Редко случается, чтобы пророчество оказалось настолько бездарным: последующие годы принесли сотни искусственных синтезов такого рода, причем во многих из них использовались основополагающие разработки Кораны. Протагонистом противоположной точки зрения явился Артур Корнберг, который сказал, обращаясь к Коране: "То, что Вы сделали, – это атомная бомба 1980 года".

Первый синтез гена явился предшественником растянутого во времени большого методического взрыва 1970–1980-х годов в молекулярной биологии. К самым ярким вехам этого взрыва, наряду с олигонуклеотидным синтезом (с переходом его от искусства к ремеслу), можно отнести эндонуклеазы рестрикции, молекулярное клонирование, секвенирование нуклеиновых кислот и полимеразную цепную реакцию. В сочетании с многими другими находками, перечислить которые невозможно, это сделало молекулярную биологию

методически почти всемогущей. Будущее покажет, достаточно ли этого, чтобы сбросить покровы с тайны живого.

Бурный поток оригинальных публикаций делает особенно важным жанр обзора в различных его видах. Книга – сложнейший из них, хотя современные возможности работы с литературой далеко отодвинули пределы осуществимого в этой области. Л.И. Патрушев поставил перед собой трудную задачу в одиночку охватить значительную часть современного массива достижений и перспектив молекулярной и клеточной биологии, непосредственно связанную с нуклеиновыми кислотами. На этом пути он сумел сделать удивительно много. Достаточно просмотреть оглавление, чтобы убедиться, что в книге не оставлены без внимания почти все основные направления, имеющие прямое отношение к нуклеиновым кислотам. При этом использован новейший материал, включая множество данных, опубликованных в этом году. Работа не прекращалась ни на минуту, и даже последние числа минувшего октября были свидетелями рождения новых разделов. В результате создано большое молекулярно-биологическое полотно, в котором каждый – от студента до (как это ни кощунственно) академика – может найти для себя что-то неожиданное.

По способу изложения материала квалифицировать книгу непросто. Ее не назовешь учебником в строгом смысле слова, хотя систематичность и внутренняя логика несомненны. Нередко от читателя требуется солидная подготовка, способность истолковывать аллюзии и проскакивать над эллипсисами. Зато удовлетворение от проникновения в материал очень велико.

Эта книга – не просто крупномасштабная компиляция. На ней отчетливо видна печать личности автора. Чувствуется, как небезразлично ему все то, что вот уже несколько десятилетий происходит на огромном пространстве молекулярной биологии. Многое ему удалось превосходно, хотя подстерегающие на этом пути тернии пристрастия, в частности несбалансированной нежности к порождениям собственного разума, не всегда увиты лаврами. В то же время, именно таков путь к звездам единой точки зрения, может быть, даже всеохватывающей концепции – в определенном смысле аналога единой теории поля, мысли о которой когда-то превратили в интеллектуальную трагедию закат великого физика.

Предлагаемая книга – плод эрудиции, вдохновения и самоотверженного труда Л.И. Патрушева и тех, чьи мысли и дела он цитирует, обсуждает, превозносит или критикует, – несомненно найдет множество благодарных читателей.

Ю. Берлин

Ноябрь 1999 г.

Предисловие автора

Мы можем понять что-то в природе, только если мы размышляем о ней...

Вернер Гейзенберг

Предметные исследования, осознанно устремляясь к границам и истокам нашего бытия, становятся философскими.

Карл Ясперс

Направления молекулярной генетики, основанные на генной инженерии, начало которым было положено в 1970-х годах работами В. Арбера, П. Берга, Г. Бойера, С. Козна и других зарубежных ученых, получили в последующие 25 лет стремительное развитие, открывшее новые горизонты биологической науки. Сегодня трудно представить себе какой-либо раздел молекулярной биологии или генетики, в котором не применялись бы достижения этих исследований. Области использования методов молекулярной генетики непрерывно расширяются, все больше охватывая, к примеру, фундаментальную и клиническую медицину, судебно-следственную практику и даже палеонтологию. Однако все эти достижения – лишь внешняя, хорошо видимая часть айсберга. Замена поврежденных генов на здоровые в генотерапии и "клонирование" многоклеточных организмов придают нашей жизни совершенно новое, грозное качество, так как ставят перед человечеством неведомые ранее этические проблемы. Для меня несомненно, что именно с 70-х годов XX в. начался новый этап эволюции биосферы Земли, все последствия которого мы еще не в состоянии предвидеть.

Работая, как и большинство из нас, в узкой области молекулярной генетики, я часто затрачивал значительные усилия на поиск сведений о каком-либо новом методе или молекулярном механизме, описание которых можно найти иногда лишь в труднодоступной англоязычной литературе. Постоянное общение с коллегами показывает, что блуждания по научной литературе в поисках ответа на теоретические вопросы, возникающие в процессе экспериментальной работы, не являются исключительно моими трудностями. Поэтому стремление сохранить время и облегчить жизнь коллегам-

экспериментаторам было одной из движущих сил, заставивших написать эту книгу.

Однако не только указанная причина побудила меня взяться за перо. Обратной стороной лавинообразного накопления научной информации является размывание контуров целостной картины современных молекулярно-генетических знаний. В соответствии с формулировкой Г. Вейля (берущей начало от Г. Галилея), метод естественных наук включает в себя пассивное наблюдение, уточняемое с помощью активного эксперимента, и построение символической конструкции, к которой в конечном счете сводятся естественно-научные теории. Высшей формой символической конструкции является математическая формула. На пути к своей цели – символической конструкции, непротиворечиво объединяющей эмпирические факты, – естественно-научные теории неизбежно проходят предварительные стадии, в том числе и стадию классификации или морфологии.

То, что в основном происходит сегодня в нашей молодой области биологии, напоминает скрупулезные усилия ботаников или зоологов-систематиков, стремящихся как можно полнее описать фауну и флору нашей планеты. Открываются все новые и новые гены, ферменты, биохимические механизмы и метаболические пути, которые переплетаются в сложнейшие сети. Миллиарды пар оснований отсеквенированных последовательностей ДНК призваны внести ясность в наше понимание структуры и функционирования генома. И хотя это совершенно закономерный и необходимый этап развития науки, одних только эмпирических фактов явно недостаточно для наступления ясности.

Складывающуюся ситуацию, в частности хорошо иллюстрирует положение в современной онкологии. Громадные усилия и средства, затрачиваемые в мире на изучение молекулярных механизмов канцерогенеза, приносят обильные плоды в виде новых генов и белков-регуляторов клеточного цикла, передачи сигнала и механизмов регуляции апоптоза. Однако прямолинейные попытки управления экспрессией этих конкретных генов с помощью антисмысловых РНК, рибозимов, анти-генов и других современных молекулярно-генетических подходов пока не привели к удовлетворительным результатам в генотерапии онкологических (и каких-либо других системных) заболеваний. Заболеваемость раком, в том числе и в промышленно развитых

странах, непрерывно возрастает и по прогнозам Всемирной организации здравоохранения к 2020 году удвоится. Пока не слишком успешными остаются и попытки борьбы со СПИДом. Такого рода факты ясно показывают, насколько далеки мы сегодня от понимания живого организма как целого.

Кроме эмпирических фактов требуются новые обобщающие идеи и в конечном счете фундаментальные символические конструкции, которые бы продвинули генетику на пути понимания законов генетического контроля жизнедеятельности клетки и организма в целом. В этой связи не менее важной целью написания книги была попытка систематизировать современные достижения в исследовании механизмов регуляции экспрессии генов на всех основных уровнях реализации генетической информации и обсудить ряд новых концепций, рожденных в ходе исследований. Я надеялся, что некоторые не слишком широко известные факты в свою очередь помогут генерировать новые знания. Однако книга не является полностью компилятивной. Наряду с современной классикой в ней представлены и результаты оригинальных исследований, которые включают, в частности концепции ксенобиоза и альтруистичной ДНК, а также гипотезы о функциональном значении редактирования РНК на посттранскрипционном уровне и образовании доминантных мутаций с участием антисмысловых РНК.

Книга состоит из двух частей. В первой части рассматриваются молекулярные механизмы, обеспечивающие хранение генетической информации живых организмов и ее реализацию через экспрессию генов. При этом основное внимание уделяется генетическим системам эукариот. Поскольку любой ген существует благодаря передаче его точной копии соматическим клеткам и в ряду поколений организмов, в конце первой части обсуждаются механизмы репликации ДНК, репарации ее повреждений, а также другие способы стабилизации генетической информации.

Вторая часть книги посвящена обсуждению современных достижений молекулярной генетики в ее прикладных областях. Однако использование термина "прикладные" в этом контексте не совсем корректно. Результаты, полученные в исследованиях трансгеноза, рибозимов или при решении задач белковой инженерии изменили облик всей современной генетики и потребовали пересмотра многих теоретических представлений.

К сожалению, ограничения в объеме книги не позволили реализовать в ней первоначально задуманную главу "Гены и фенотип", а также обсудить современные достижения в исследовании рекомбинации ДНК. По тем же причинам упомянуты не все авторы, работы которых обсуждаются в монографии. Им я приношу свои глубокие извинения. Исчерпывающую библиографию можно найти в современных обзорах из списка рекомендованной литературы, которая предназначена для более глубокого знакомства с предметом.

На протяжении всей книги мне хотелось подчеркнуть единство генетических систем организма. Отдельное рассмотрение систем транскрипции, трансляции, репарации или репликации условно, поскольку их функционирование в клетке тесно взаимосвязано. На это указывает, в частности высокая интеграция многочисленных механизмов внутриклеточной передачи сигнала, полифункциональность отдельных белков и ферментов, колоссальное разнообразие летальных мутаций и плейотропность действия генов.

Работая над книгой, я всегда имел в виду, что ясное представление о целом может быть хорошим помощником в исследовании частных генетических проблем. Но и сама генетика является лишь малой частью творения духа человека. В лаборатории, клинике, на пленэре или в мастерской мы пишем этюды к большой картине, которой, к счастью, никогда не суждено быть завершенной. И в те редкие минуты, когда работа идет особенно хорошо, отчетливо ощущаешь причастность к этому великому движению духовных сфер, во многом определяющему смысл нашей жизни.

В заключение предисловия хочется выразить особую признательность дирекции и Ученому совету Института биоорганической химии им. М.М. Шемякина и Ю.А. Овчинникова РАН за оказанную поддержку в издании книги.

Работа над книгой проходила параллельно с экспериментальными исследованиями в стенах лаборатории биотехнологии ИБХ, руководимой А.И. Мирошниковым, в лице которого я нашел полное понимание. За это я хочу принести ему свою искреннюю благодарность.

Мое мировоззрение как генетика начало складываться и во многом определилось в аспирантуре на семинарах Р.Б. Хесина, а также при личном общении с этим замечательным человеком и сотрудниками его лаборатории. Ко всем этим людям, ученым и учителям, с которыми мне посчастливилось работать в начале жизненного пути, я храню чувство глубокой благодарности.

Я отдаю себе отчет в том, что монография никогда не была бы написана без высокого профессионализма, постоянного интереса и дружеского участия заведующей научно-информационным отделом Института Т.И. Соркиной, от которой исходит сама идея этого издания. В долгих и плодотворных дискуссиях с ней обсуждалась общая структура книги и содержание отдельных глав, а также многочисленные технические детали, о которых перед началом работы я, к сожалению, не имел ни малейшего представления. Одновременно я приношу свою глубокую благодарность сотрудникам отдела В.В. Егоровой, Т.И. Яковлевой, И.М. Приваловой и Л.И. Петровой за большую техническую помощь при подготовке рукописи к печати.

Хотелось бы выразить искреннюю признательность Е.Д. Свердлову за жесткую и конструктивную критику первых двух глав монографии, которая оказала сильное влияние на всю дальнейшую работу над книгой. Сейчас невозможно представить себе выход монографии в свет без учета многочисленных глубоких замечаний Ю.А. Берлина, у которого я многому научился в ходе нашего продолжительного общения.

Пользуясь случаем, хотелось бы выразить особую признательность всем зарубежным коллегам, приславшим оттки свои недавних публикаций.

И в заключение, но не в последнюю очередь, я хотел бы от всей души поблагодарить свою жену Раю, такт, мудрость и душевная аура которой, несмотря на все мои недостатки, помогают поддерживать творческую атмосферу в семье. Без всего этого написание книги было бы невозможно.

Неоценимая помощь профессионалов не снимает с меня ответственности за фактическую сторону материала, представленного в монографии. Я буду благодарен за любую конструктивную критику всех спорных интерпретаций генетических фактов, а также замеченных ошибок и неточностей, которые, к сожалению, чаще всего можно увидеть только на расстоянии.

Л.И. Патрушев

ЧАСТЬ I. МЕХАНИЗМЫ ХРАНЕНИЯ И РЕАЛИЗАЦИИ ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ

ВВЕДЕНИЕ

Организм. Живой организм представляет собой самовоспроизводящуюся, открытую термодинамическую систему, в которой пути превращения вещества и энергии определяются генетической информацией, реализующейся через генетический код. В таком несколько измененном определении М. Барбьери (1998 г.) сформулирован один из фундаментальных принципов биологии конца XX в. Сама возможность появления этого глобального обобщения, по мнению Барбьери, связана с тремя крупнейшими прорывами в биологических исследованиях уходящего века. Биохимиками была определена функциональная роль белков, генетикам удалось приподнять завесу над миром генетической информации, а молекулярные биологи установили, что связь между этими двумя мирами осуществляется через генетический код.

Поскольку организм является самовоспроизводящейся системой, он должен обладать информацией, необходимой и достаточной для поддержания своего *status quo* и производства себе подобных. Словосочетание *открытая термодинамическая система* подчеркивает, что все проявления жизнедеятельности организма могут быть объяснены в терминах естественных наук, а одним из необходимых условий его существования является обмен с окружающей средой веществом, энергией и информацией в разных формах, включая генетическую информацию.

По современным представлениям вся генетическая информация живого организма содержится в его генах и, как любая другая информация, заключает в себе *сообщения*, в данном случае сообщения для молекулярных объектов, способных их воспринять. Сама возможность восприятия генетической информации определяется тем, что сообщения организованы с помощью системы правил (*генетического кода*), "понятных" объектам, для которых эти сообщения предназначены. Получив генетическое сообщение, молекулярный объект его декодирует в соответствии с правилами, лежащими в основе функционирования его составных частей. Даже частичное невыполнение этих правил может приводить к тяжелым нарушениям жизнедеятельности организма.

Во время передачи генетической информации по существующим каналам связи от генов к воспринимающим молекулярным объектам имеет место ее многократное декодирование и перекодирование вплоть до окончательного воплощения в фенотипических признаках. Происходит *экспрессия генов*. Можно сказать, что в результате экспрессии генов заключенное в них слово материализуется в деяние. Однако это не более чем метафора: природа без человека лишена членораздельной речи. Было бы непродуктивно, не опираясь на демиурга, искать в строении организмов предсуществующий план, а в действии генов – цель.

Клетка. Клетка является той наименьшей частью организованной материи, которая еще сохраняет все признаки живого. В ней происходят основные события, связанные с экспрессией генов. Лишь в клетке может полноценно реализовываться генетическая информация, заключенная в генах. Рассмотрение генов в отрыве от их естественного внутриклеточного окружения приводит к искусственному разделению генома на функционально оторванные друг от друга фрагменты, созданию емких баз данных последовательностей нуклеиновых кислот. Современные попытки заставить функционировать гены в новом генетическом окружении завершаются в лучшем случае получением нежизнеспособных в природных условиях биореакторов – организмов-сверхпродуцентов собственных или чужеродных белков и метаболитов.

Связь между генами и внутриклеточным генетическим окружением неразрывна. Компоненты клетки, распознавая гены, считывают заключенную в них информацию и декодируют ее. При этом они поддерживают гены в рабочем состоянии и их воспроизводят, создавая точную копию носителя информации, гарантию существования самих себя. Действительно, в ответ на это гены предоставляют инструкции, необходимые для внутриклеточного клонирования самих молекулярных компонентов клеток, а следовательно (по крайней мере, в случае одноклеточных организмов) создания их точных копий. Функционирует единая система, части которой невозможно разъединить, не уничтожив живое. Тем не менее, суть связей между компонентами генетической системы, организованной в клетку, – это не порочный круг. Напротив, рассматривая систему, как единое целое, в ней можно обнаружить внутренний потенциал направленного развития, который реализуется во времени, материализуясь в биосфере.

В этой связи популярные еще и сегодня рассуждения об эгоистическом гене, заставляющем клетку работать только на себя, лишены глубокого смысла. Последовательно оставаясь на такого рода антропоморфных позициях, необходимо признать, что и эгоистически настроенные белки могут рассматривать ДНК как сожителя, существование которого приходится терпеть для воспроизводства самих себя.

"Omnia mea mecum porto!"¹ – этот девиз в полной мере применим к одноклеточным организмам. У многоклеточных генетических систем это не так. В данном случае отдельные клетки, составляющие организм, становятся зависимыми друг от друга, создавая неразрывное единство на другом уровне – сомы, или тела. Между клетками происходит контролируемое генами перераспределение функций. При этом клетки разных частей многоклеточного организма могут настолько различаться морфологически, что без проведения специального анализа на молекулярном уровне их невозможно отнести к одной генетической системе, единому организму.

Пролиферация, дифференцировка и апоптоз соматических клеток.

Переход к многоклеточности перевел отношения организма с окружающим миром на новый уровень сложности. Многочисленные связи между рецепторами организма и внешними воздействиями, с одной стороны, а также внутренними анализаторами поступающей информации, с другой, обеспечивают максимальную приспособленность организма к условиям существования. Жизнь не прощает ошибок. Все, что неадекватно реагирует на сигналы среды существования, элиминируется естественным отбором. Чем чувствительнее и эффективнее система настроена на окружающий мир (в том числе и внутреннюю среду организма), тем сложнее и изощреннее ее морфологическое воплощение, ее внутреннее содержание.

Пролиферация клеток. В основе развития любого многоклеточного организма, становления системы его органов и тканей лежит деление (пролиферация) клеток. Генетическая программа обеспечивает протекание сложной совокупности биохимических реакций, сопровождаемых созданием точной копии генетического аппарата каждой соматической клетки, ее ростом и делением. Поскольку при каждом делении клеток весь глобальный

¹ Все свое ношу с собой (лат.)

биохимический процесс циклически повторяется, он получил название *клеточного цикла*. Индивидуальное (онтогенетическое) развитие, как правило, начинается с первого деления стимулированного к этому яйца (яйцеклетки) и завершается только с наступлением его смерти – распада организма как целого в результате обрыва ключевых внутренних связей между системами его жизнеобеспечения. Основное внутреннее событие жизни организма – деление клеток – находится под строгим внутренним и внешним генетическим контролем. Даже изолированная соматическая клетка способна лишь к ограниченному количеству делений в питательной среде. Количественный контроль числа клеточных делений лежит в основе органогенеза (формирования органов и тканей). Нарушение механизмов контроля пролиферации клеток приводит к безудержному делению клеток, образованию бесформенной клеточной массы – опухоли, способной задушить организм изнутри. Однако в процессе нормального онтогенетического развития изменяется не только число соматических клеток, но и их качественный состав.

Дифференцировка клеток. Способность органов и тканей осуществлять свои специфические функции целиком зависит от наличия в них специализированных клеток. В частности, организм взрослого человека составлен из $\sim 10^{14}$ – 10^{15} клеток более чем 100 различных типов. На самых ранних стадиях развития зародыша многоклеточного организма составляющие его клетки внешне очень похожи друг на друга. По мере продолжения онтогенеза пути многих из них далеко расходятся. Происходит дифференцировка клеток, приобретение ими специализированных функций. Морфологические различия, выявляемые у специализированных клеток, определяются особым составом и внутриклеточной организацией их молекул. Появление таких особенностей на молекулярном уровне также контролируется генами. В специализированных (дифференцированных) клетках или их предшественниках кроме генов, экспрессирующихся в клетках всех типов, работают особые группы генов. Переключение экспрессии одних групп генов на другие, подключение к экспрессии новых генов и прекращение работы старых в дифференцирующихся клетках также находится под строгим генетическим контролем.

Апоптоз. Еще одним важным событием индивидуального развития организма является полное замещение одних групп клеток другими. При этом

конечные стадии процесса замены контролируются самими замещаемыми клетками. На определенной стадии развития эмбриона в ответ на сигналы окружающих тканей внутри удаляемых клеток происходит активация группы генов, запускающих их саморазрушение – апоптоз. Апоптоз является одним из проявлений принципа самоочищения организма, когда для становления и сохранения целого многоклеточный организм жертвует небольшой частью своих соматических клеток. Действительно, другой не менее важной стороной этого процесса является защита организма от клеток с необратимо поврежденным генетическим аппаратом, поскольку в этом случае возникает опасность их неконтролируемого роста и гибели целого организма. В том случае, если повреждения генетического аппарата клетки невозможно восстановить, клетка совершает самоубийство.

С учетом всего сказанного можно без преувеличения утверждать, что пролиферация, дифференцировка и апоптоз соматических клеток определяют ключевые моменты внутренней жизни многоклеточного организма.

Гены. В соответствии с центральным постулатом молекулярной биологии принято считать, что генетическая информация, необходимая для индивидуального развития организма, заключена в его генах, которые представляют собой последовательности нуклеотидов молекул ДНК и РНК. Гены содержат информацию о составных частях организма: совокупности большого числа высоко- и низкомолекулярных химических соединений, образующих его клетки, ткани и органы. При этом данные о структуре почти всех низкомолекулярных соединений, называемых *метаболитами*, закодированы в генах не прямо, а косвенно. Фактически эта информация является лишь программой биосинтеза метаболитов. Сама же структура и взаимодействие метаболитов друг с другом определяются биологическими катализаторами белковой природы – ферментами, которые и осуществляют необходимые их взаимопревращения – *метаболизм*.

Генетическая информация о структуре белков и ферментов скрыта в генах не так глубоко. Благодаря существованию универсального триплетного генетического кода, последовательности нуклеотидов генов однозначно определяют последовательности аминокислот полипептидных цепей конкретных белков. Декодирование информации о структуре белков и нуклеиновых кислот, сопровождаемое их биосинтезом, является важнейшим

промежуточным (но не конечным) результатом функционирования (экспрессии) генов любого организма.

Экспрессия генов. В норме экспрессия генов обеспечивает существование организма как целого от начальных до завершающих стадий индивидуального развития – от первых делений стимулированной яйцеклетки до естественной смерти организма. Однако более широкий подход к проблеме экспрессии генов должен учитывать не только биохимические последствия их работы на молекулярном, надмолекулярном и организменном уровне, но и генетически детерминированные поведенческие реакции групп особей в популяции, а следовательно, и механизмы генетического контроля развития самих популяций, включая цивилизацию.

При реализации запрограммированных фенотипических эффектов работающие вместе гены и продукты их экспрессии правильно декодируют адресованные им (и только им) регуляторные сообщения на всех уровнях и адекватно на них отвечают. Адресная доставка и расшифровка как регуляторных сигналов, так и самой генетической информации, заключенной в генах, становятся возможными благодаря осуществлению высокоспецифических молекулярных взаимодействий.

Специфичность всех биохимических реакций обеспечивается способностью высокомолекулярных соединений организма безошибочно распознавать друг друга и низкомолекулярные метаболиты. Специфическое межмолекулярное узнавание, которое определяет упорядоченное протекание всех генетических процессов, осуществляется, по крайней мере, на трех уровнях. На первом уровне (в соответствии с последовательностью основных событий, происходящих при реализации генетической информации) имеет место специфическое распознавание друг друга нуклеиновыми кислотами. Цепи одной молекулы ДНК взаимодействуют между собой по принципу комплементарности в соответствии с правилами Уотсона–Крика, что характерно и для соответствующих контактов ДНК с РНК, а также для взаимодействия молекул РНК друг с другом. Такой тип взаимодействий лежит в основе самовоспроизведения генетической информации и ее передачи от нуклеиновых кислот к нуклеиновым кислотам или белкам. Второй уровень обеспечивается белково–нуклеиновым узнаванием. Этот крайне важный тип взаимодействий регулирует экспрессию генов и способствует упорядоченному

метаболизму нуклеиновых кислот. И, наконец, взаимодействие белков друг с другом, с иными макромолекулами клеток или с низкомолекулярными лигандами делает возможной сборку молекулярных и надмолекулярных комплексов, обеспечивающих направленное протекание метаболических процессов, и лежит в основе морфогенеза организмов. Кроме всего прочего, этот тип взаимодействий может изменять специфичность действия белков и индуцировать в них новую активность.

Постепенно становится ясно, что во время реализации генетической информации в процессе биосинтеза нуклеиновых кислот, белков и ферментов формируются сложные пространственные структуры из белков и нуклеиновых кислот в результате их самоорганизации, часто при участии других молекул клетки. Самоорганизация синтезированных макромолекул приводит к образованию пространственных внутриклеточных межмолекулярных структур, клеточных органелл и, в конечном счете, самих клеток, которые в совокупности создают многоклеточный организм, законы функционирования и предназначение которого во многом остаются непонятными и сегодня.

ГЛАВА 1. ГЕНОМ

Термин "геном" был предложен Г. Винклером в 1920 г. для описания совокупности генов, заключенных в гаплоидном наборе хромосом организмов одного биологического вида. Первоначальный смысл этого термина указывал на то, что понятие генома в отличие от генотипа является генетической характеристикой вида в целом, а не отдельной особи. С развитием молекулярной генетики значение данного термина изменилось. Сейчас хорошо известно, что ДНК, которая является носителем генетической информации у большинства организмов и, следовательно, составляет основу генома, включает в себя не только гены в современном смысле этого слова. Большая часть ДНК эукариотических клеток представлена некодирующими ("избыточными") последовательностями нуклеотидов, которые не заключают в себе информации о белках и нуклеиновых кислотах. **Таким образом, основную часть генома любого организма составляет вся ДНК его гаплоидного набора хромосом.**

Однако генетическую информацию в клетках содержат не только хромосомы ядра. Жизненно важная генетическая информация бывает заключена и во внехромосомных молекулах ДНК. У бактерий к таким ДНК относятся плазмиды и некоторые умеренные вирусы, в клетках эукариот – это ДНК хлоропластов, митохондрий и других пластид. Более того, объемы генетической информации, заключенной в клетках зародышевой линии (предшественники половых клеток и сами гаметы) и соматических клетках, в ряде случаев существенно различаются. В онтогенезе соматические клетки могут утрачивать часть генетической информации клеток зародышевой линии, амплифицировать группы последовательностей и(или) значительно перестраивать исходные гены. Следовательно, под геномом организма в настоящее время понимают суммарную ДНК гаплоидного набора хромосом и каждого из внехромосомных генетических элементов, содержащуюся в отдельной клетке зародышевой линии многоклеточного организма. Однако сформулировать определение генома отдельного биологического вида в целом не так просто. В таком определении необходимо учитывать, во-первых, генетические различия, связанные с полом организма, поскольку мужские и

женские половые хромосомы различаются. Во-вторых, из-за громадного числа аллельных вариантов генов и сопутствующих последовательностей, которые присутствуют в генофонде больших популяций, можно говорить лишь о некоем усредненном геноме, который сам по себе может обладать существенными отличиями от геномов отдельных особей.

Как видно из табл. I.1, размеры геномов организмов разных видов значительно отличаются друг от друга. При этом часто не наблюдается корреляции между уровнем эволюционной сложности биологического вида и размером его генома.

Таблица I.1

Средний размер гаплоидного генома у некоторых групп организмов

Группы организмов	Средний размер генома, п.о.
Мелкие вирусы	$1,0 \cdot 10^4$
Микоплазмы	$1,6 \cdot 10^6$
Бактерии	$2,0 \cdot 10^6$
Грибы	$4,7 \cdot 10^7$
Насекомые	$2,3 \cdot 10^9$
Моллюски	$1,6 \cdot 10^9$
Костистые рыбы	$1,4 \cdot 10^9$
Амфибии бесхвостые	$2,7 \cdot 10^9$
хвостатые	$3,6 \cdot 10^{10}$
Рептилии	$1,5 \cdot 10^9$
Птицы	$1,2 \cdot 10^9$
Млекопитающие	$2,6 \cdot 10^9$
человек	$3,0 \cdot 10^9$
Растения голосеменные	$1,6 \cdot 10^{10}$
покрытосеменные	$2,7 \cdot 10^{10}$
лилия <i>Lilium longiflorum</i>	$1,8 \cdot 10^{11}$

Суммарное количество ДНК в гаплоидном геноме принято обозначать латинским символом *C*. В 1978 г. Т. Кавалье-Смит описал в качестве парадокса наблюдение того, что у эукариот транскрибируется лишь незначительная часть последовательностей нуклеотидов генома (**~3% генома человека**). В соответствии с этим необъясненный до недавнего времени феномен значительной избыточности генома эукариот в отношении некодирующих последовательностей нуклеотидов известен в генетике под названием **"парадокса C"**. По моему мнению (см. раздел 5.3), эволюционное включение избыточных последовательностей нуклеотидов в исходный геном-предшественник стабилизировало заключенную в нем генетическую информацию, что, в свою очередь, создало необходимые условия для возникновения многоклеточности в природе.

Структурная организация генома является фундаментальным таксономическим признаком, лежащим в основе современной систематики животного и растительного мира. **В соответствии со структурной организацией генома все живые организмы разделяют на два надцарства: прокариот и эукариот.** К прокариотам относят организмы, геном которых не заключен в ядро, ограниченное ядерной мембраной, и его редупликация не сопровождается митозом. Надцарство прокариот включает в себя три царства: **архебактерий (Archaeobacteria), шизомикофитов (Shizomycophyta) и цианофитов (Cyanophyta).** Таким образом, к прокариотам относятся сине-зеленые водоросли, актиномицеты, все бактерии, микоплазмы, риккетсии и вирусы.

Клетки эукариот содержат оформленное ядро, и редупликация их генома сопровождается митозом. **В надцарство эукариот входят царства: мезокариоты (жгутиконосцы) (Mesokaryotes), грибы (Fungi), растения (Planta) и животные (Animalia).** Характерная структура генома прокариот и эукариот накладывает отпечаток на морфологические, физиологические, биохимические и генетические особенности этих организмов, которые в конечном счете определяются генетической информацией, заключенной в их геномах и реализующейся через экспрессию соответствующих генов. Таким образом, можно без преувеличения сказать, что структура генома лежит в основе всех тех внутренних и внешних проявлений жизнедеятельности любого организма, которые определяют его положение в иерархии живых существ,

населяющих нашу планету.

1.1. Гены и хромосомы

Генетическая информация о структуре отдельных белков и нуклеиновых кислот у всех организмов заключена в молекулах ДНК или РНК в виде последовательностей нуклеотидов, называемых **генами**. Однако одной информации о структуре макромолекул, кодируемых генами, недостаточно для их функционирования. Координированная работа (экспрессия) большого числа генов возможна лишь благодаря наличию тонких регуляторных механизмов, определяющих место, время и уровень экспрессии конкретного гена или группы генов. Для того чтобы экспрессия гена была регулируемой, он должен содержать индивидуальную (регуляторную) метку, по которой регуляторные компоненты генетической системы клетки или организма могли бы безошибочно оказать на него необходимое воздействие. В соответствии с этим **любой ген состоит из двух основных функциональных частей (последовательностей нуклеотидов) – регуляторной и структурной**. Регуляторная часть обеспечивает первые этапы реализации генетической информации, заключенной в структурной части гена, которая, в свою очередь, содержит информацию о структуре конкретных белков или нуклеиновых кислот. Поэтому размер гена складывается из размеров его структурной и регуляторной частей. Однако определить протяженность гена не так просто, особенно в случае генов эукариот.

Отдельные элементы регуляторной области генов, например энхансеры, могут располагаться на значительном (>60 т.п.о.) расстоянии от структурной части гена как перед ней, так и позади нее или даже в ней самой. В самой же структурной части большинства эукариотических генов кодирующие последовательности нуклеотидов (**экзоны**) перемежаются протяженными некодирующими последовательностями (**интронами**). Суммарный размер интронов, как правило, многократно превышает суммарный размер экзонов конкретных генов. Уже исходя из этого факта, можно сделать вывод о том, что геном любого эукариотического организма содержит не только последовательность нуклеотидов с генетической информацией о белках и нуклеиновых кислотах, но и большое количество последовательностей нуклеотидов, не несущих такой информации.

Однако помимо интронов в геноме эукариот имеется большое количество других некодирующих последовательностей нуклеотидов, главным образом различных повторяющихся последовательностей. Поэтому общая длина некодирующих последовательностей нуклеотидов в геноме эукариот в десятки раз превышает длину кодирующих последовательностей. Не вполне определенные и очень большие размеры генов эукариот, к тому же расположенных в геноме среди многочисленных некодирующих последовательностей нуклеотидов, создают значительные трудности для изучения их структуры и функционирования *in vivo*.

Как у прокариотических, так и у эукариотических организмов все гены располагаются группами на отдельных молекулах ДНК, которые при участии белков и других макромолекул клеток организуются в хромосомы. Зрелые клетки зародышевой линии (*гаметы* – яйцеклетки, спермии) многоклеточных организмов содержат по одному (*гаплоидному*) набору хромосом организма. У *диплоидных* (*полиплоидных*) организмов, клетки которых содержат по одному (несколько) набору хромосом каждого из родителей, одинаковые хромосомы получили название **гомологичных хромосом**, или **гомологов**. Гомологичными являются и одинаковые хромосомы разных организмов одного биологического вида. Гены и некодирующие последовательности нуклеотидов, заключенные в хромосомах ядер клеток, представляют большую часть генома организма. Кроме того, геном организма формируют и *внехромосомные генетические элементы*, которые во время митотического цикла воспроизводятся независимо от хромосом ядер. Так, митохондрии грибов и млекопитающих содержат менее 1% всей ДНК, тогда как у почкующихся дрожжей *Sacharomyces cerevisiae* митохондриальная ДНК составляет до 20% всей ДНК клетки. ДНК пластид растений (главным образом хлоропластов и митохондрий) составляет от 1 до 10% суммарного количества их ДНК.

Поскольку гены, входящие в состав отдельных хромосом, находятся в одной молекуле ДНК, они образуют отдельную **группу сцепления** и в отсутствие рекомбинации вместе передаются от родительских клеток к дочерним. Остаются до конца не понятыми физиологическое значение распределения генов по отдельным хромосомам и природа факторов, определяющих число хромосом в геноме эукариот. Например, невозможно объяснить эволюционные механизмы появления большого числа хромосом у

конкретных организмов только ограничениями, накладываемыми на максимальный размер молекул ДНК, входящих в состав этих хромосом. Так, геном американской амфибии *Amphiuma* содержит в ~30 раз больше ДНК, чем геном человека, и вся ДНК заключена только в 28 хромосомах, что вполне сопоставимо с кариотипом человека (46 хромосом). Однако даже самая маленькая из этих хромосом больше самых крупных хромосом человека. Остаются неизвестными и факторы, ограничивающие верхний предел числа хромосом у эукариот. Например, у бабочки *Lysandra nivescens* диплоидный набор составляет 380–382 хромосомы, и нет основания считать, что это значение является максимально возможным.

По-видимому, большинство особенностей структурной и функциональной организации генома должны обеспечивать надежность его функционирования, т.е. точность передачи генетической информации от родительских клеток дочерним на протяжении многих клеточных генераций и правильную работу генов. Поэтому можно предполагать, что они имеют, прежде всего, отношение к обеспечению надежности передачи и внутриклеточной реализации генетической информации, проявляющейся в упорядоченной во времени и безошибочной экспрессии генов.

1.2. Геном прокариот

Как уже было упомянуто выше, основной чертой молекулярной организации прокариот является отсутствие в их клетках (или *вирионах* – вирусных частицах, в случае вирусов) ядра, отгороженного ядерной мембраной от цитоплазмы, если она существует. Отсутствие ядра является лишь внешним проявлением особой организации генома у прокариот, которая коренным образом отличается от таковой у эукариотических организмов. В отличие от эукариот, геном прокариот построен очень компактно. Количество не кодирующих последовательностей нуклеотидов минимально, интроны редки. Более того, у прокариот для кодирования белков часто используются две или все три рамки считывания одной и той же последовательности нуклеотидов гена, что повышает кодирующий потенциал их генома без увеличения его размера. Многие механизмы регуляции экспрессии генов, использующиеся у эукариот, никогда не встречаются у прокариот. Это не относится к вирусам животных и растений, которые, являясь внутриклеточными паразитами

эукариотических клеток, используют необходимую часть их генетического потенциала для своих нужд. Таким образом, **простота строения генома прокариот объясняется, прежде всего, их упрощенным жизненным циклом, на протяжении которого прокариотические клетки, как правило, не претерпевают сложных дифференцировок, связанных с глобальным переключением экспрессии одних групп генов на другие, или тонким изменением уровней их экспрессии, что имеет место в онтогенетическом развитии эукариот.**

1.2.1. Геном вирусов

По определению Х. Френкель-Конрата, "вирусы – это частицы, состоящие из одной или нескольких молекул ДНК или РНК, обычно (но не всегда) окруженных белковой оболочкой; вирусы способны передавать свои нуклеиновые кислоты от одной клетки-хозяина к другой и использовать ее ферментативный аппарат для осуществления своей внутриклеточной репликации путем наложения собственной информации на информацию клетки-хозяина; иногда вирусы могут обратимо включать свой геном в геном хозяина (интеграция), и тогда они либо ведут "скрытое существование", либо так или иначе трансформируют свойства клетки-хозяина"². В приведенном определении отмечены характерные особенности жизненного цикла вирусов, которые находят отражение в организации их генома. Вирусы являются внутриклеточными паразитами и используют для своего размножения белоксинтезирующий аппарат клетки-хозяина. Жизненный цикл вируса начинается с проникновения внутрь клетки. Для этого он связывается со специфическими рецепторами на ее поверхности и либо вводит свою нуклеиновую кислоту внутрь клетки, оставляя белки вириона на ее поверхности, либо проникает целиком в результате эндоцитоза. В последнем случае после проникновения вируса внутрь клетки следует его раздевание – освобождение геномных нуклеиновых кислот от белков оболочки, что делает вирусный геном доступным для ферментных систем клетки, обеспечивающих экспрессию генов вируса.

После проникновения вируса в клетку может происходить его

² Френкель-Конрат Х. Химия и биология вирусов. М.: Мир, 1972. С. 15.

размножение, часто сопровождаемое гибелью самой клетки (**вирулентный путь развития**). Кроме того, вирус может длительное время существовать внутри клетки, внешне ничем себя не проявляя (**латентная инфекция**). В этом случае его геном встраивается в геном клетки-хозяина и реплицируется вместе с ним или находится во внехромосомном состоянии. После проникновения вирусной геномной нуклеиновой кислоты в клетку заключенная в ней генетическая информация должна быть расшифрована генетическими системами хозяина и использована для синтеза компонентов вирусных частиц. Поскольку для своего размножения вирусы используют главным образом ферментные системы клетки-хозяина, их геном характеризуется относительно малыми размерами и кодирует структурные белки вирионов, а также белки и ферменты, которые перестраивают метаболизм клетки для нужд размножения вируса, делая процесс репликации вирусов максимально эффективным. Геном вирусов, заключенный внутри вирионов, может быть представлен одноцепочечными или двухцепочечными ДНК или РНК. Кроме того, все гены вирусов могут быть заключены в одной хромосоме или разделены на несколько блоков (хромосом), которые все вместе и составляют геном таких вирусов. Например, у реовирусов геном представлен двухцепочечной РНК и состоит из десяти сегментов. Геномы вирусов, содержащих одноцепочечную РНК, также могут быть либо цельными (например у ретровирусов), либо сегментированными (например у ортомиксовирусов или аренавирусов). Геном РНК-содержащих вирусов представлен только линейными молекулами РНК.

Все известные ДНК-содержащие вирусы позвоночных имеют геном, заключенный в одной хромосоме, линейной или кольцевой, одно- или двухцепочечной. У некоторых вирусов, например у вируса гепатита В, геном представлен кольцевой ковалентно замкнутой молекулой двухцепочечной ДНК, в обеих цепях которой в разных местах обнаружены одноцепочечные участки. У нескольких родов, например адено-ассоциированных вирусов комплементарные цепи ДНК находятся в различных вирусных частицах.

1.2.2. Нуклеоид бактериальной клетки

Каждый, кому приходилось разрушать бактериальные клетки в мягких условиях, например с помощью лизоцима или детергентов, наблюдал



Рис. 1.1. Нуклеоид *E. coli*

а – электронно-микроскопические фотографии срезов бактериальных клеток, полученные методом криофиксации. 1 и 2 – один и тот же снимок (последний ретуширован). На фотографии 2 белыми пятнами отмечены области цитоплазмы, свободные от рибосом. На фотографии 3 видны молекулы ДНК, специфически окрашенные с помощью антител; *б* – модель нуклеоида в функционально-активном состоянии А. Райтера и А. Чанга. Изображены многочисленные петли активно транскрибируемой ДНК

замечательную картину превращения легко подвижной суспензии бактериальных клеток в вязкую желеобразную массу, простое перемешивание

которой требует усилий. Это происходит из-за того, что компактно упакованные гигантские хромосомы бактериальных клеток (длина хромосомной ДНК *E. coli* составляет ~4,6 млн. п.о.) после разрушения оболочки клеток выходят в окружающую среду и свободно в ней распределяются. В лизатах бактериальных клеток их ДНК прочно ассоциированы с белками, освобождение от которых требует проведения многократных фенольных депротеинизаций. Такой простой опыт наглядно указывает на то, что в бактериальных клетках их единственная хромосома сильно компактизована и, возможно, пространственно упорядочена.

Электронно-микроскопическое изучение срезов бактериальных клеток в разных условиях и более ранние исследования бактерий с помощью светового микроскопа продемонстрировали компактное распределение ДНК в бактериальной клетке. Поскольку такие структуры отдаленно напоминали ядра эукариот, они получили название **нуклеоидов**, или ДНК-плазмы. Эти термины подчеркивают генетические функции нуклеоида, но в то же время и существенные морфологические отличия от обычных интерфазных ядер эукариот, прежде всего, отсутствие ядерной оболочки, которая бы отделяла гены бактерии от окружающей их цитоплазмы. Исследование бактериальных клеток с помощью электронной микроскопии в мягких условиях без предварительной химической фиксации показало, что нуклеоиды представлены в виде диффузно окрашенных областей, свободных от рибосом (рис. 1.1,а). При этом вытянутые участки ДНК на внешней части нуклеоидов направлены в окружающую цитоплазму. С помощью специфических антител установлено, что молекулы РНК-полимеразы, ДНК-топоизомеразы I и гистонподобного белка HU ассоциированы с нуклеоидами. Вытянутые участки ДНК по периферии нуклеоидов обычно интерпретируют как сегменты бактериальной хромосомы, вовлеченные в транскрипцию. Полагают, что эти участки состоят из петель ДНК бактериальной хромосомы, которые в зависимости от физиологического состояния клетки находятся в транскрипционно-активном состоянии или втягиваются внутрь нуклеоидов при подавлении транскрипции. Модель функционально-активного нуклеоида А.Райтера и А.Чанга представлена на рис. 1.1,б. По мнению авторов, размытая структура поверхности нуклеоидов, видимая под электронным микроскопом, отражает подвижное состояние активно транскрибируемых петель ДНК. В этой модели четко прослеживается

аналогия со структурой хромосом типа ламповых щеток у животных.

Таким образом, нуклеоид бактериальных клеток не является статическим внутриклеточным образованием или компартментом, которые можно четко определять морфологически. Напротив, во время различных фаз роста бактериальных клеток нуклеоид непрерывно меняет форму, что, по-видимому, сопряжено с транскрипционной активностью определенных бактериальных генов. Так же как и в хромосомах эукариот, ДНК нуклеоида ассоциирована со многими ДНК-связывающими белками, в частности гистоноподобными белками HU, H-NS и IHF, а также топоизомеразами, которые оказывают большое влияние на функционирование бактериальных хромосом и их внутриклеточную компактизацию. Однако детальные молекулярные механизмы конденсации бактериальной ДНК с образованием лабильных "**компактосом**" (по аналогии со стабильными нуклеосомами эукариот) пока неизвестны. В последнее время возрастает интерес к бактериальному так называемому **LP-хроматину** (low protein chromatin), для которого характерно относительно низкое содержание белкового компонента. Аналогичный LP-хроматин обнаруживают у вирусов, в митохондриях, пластидах и у динофлагеллят (жгутиконосцев). Следовательно, этот тип структурной организации генетического материала претендует на универсальность и ассоциирован с определенными формами регуляции экспрессии генов, свойственными прокариотическим организмам.

В последние годы наблюдается прогресс в исследовании первичной структуры бактериальных хромосом. Определена полная последовательность нуклеотидов хромосом паразитических бактерий: микоплазмы *Mycoplasma genitalium* и *Haemophilus influenzae*. В 1997 г. усилиями интернационального коллектива ученых была определена полная первичная структура хромосом *E. coli* и *Bacillus subtilis* длиной в ~4,6 и 4,2 млн п.о. соответственно. Все это позволяет надеяться, что в ближайшее время произойдут новые открытия в области исследований структуры бактериальных геномов и функционирования их генов.

1.2.3. Геном архебактерий

Царство архебактерий представляет собой своеобразную и наименее изученную таксономическую группу прокариот. Хотя по своей морфологии

Archeabacteria похожи на привычные эубактерии, на молекулярном уровне они сближены с эукариотами. Эти микроорганизмы часто рассматривают как прокариотические эволюционные предшественники эукариот, в связи с чем представляется целесообразным рассмотреть строение генома архебактерий более подробно.

Архебактерия *Methanococcus jannaschii*, первичная структура генома которой была полностью определена в 1996 г., обнаружена в горячих морских глубоководных источниках. Энергию для жизнедеятельности этот микроорганизм получает при восстановлении двуокиси углерода до метана молекулярным водородом. Температура, близкая к температуре кипящей воды, является оптимальной для его роста, который может происходить при давлении более 200 атм. *M. jannaschii* не требует для своего роста органических соединений: все необходимое для жизни он синтезирует из неорганических веществ – CO₂, NH₃ и т.п. Геном *M. jannaschii* состоит из основной кольцевой хромосомы и двух небольших внехромосомных элементов, размеры которых составляют соответственно 1700, 58 и 16 т.п.о. Подобные размеры геномов типичны для архе- и эубактерий. Интересно, что **GC-состав ДНК этого ярко выраженного термофила невысок и составляет всего 31%**. Геном организован компактно: обнаружено **~1700 потенциальных кодирующих участков ДНК, по одному на каждые 1000 п.о.**

Многие ДНК-локусы *M. jannaschii* не обнаруживают гомологии с уже известными последовательностями. Функциональное значение большого числа потенциальных кодирующих последовательностей генома этого микроорганизма остается невыясненным. Таким образом, *M. jannaschii* отличается от других прокариот и эукариот большим набором только ему свойственных генов и функций. Анализ структуры генома *M. jannaschii* показал, что гены, организующие системы обработки генетической информации – транскрипции, трансляции и репликации ДНК, в большей степени напоминают гены эукариот, чем бактерий. При этом гены системы трансляции оказались наиболее консервативными (обладали наибольшей гомологией) у прокариот, эукариот и архебактерий. Из них гены рРНК – универсальны, так же как и гены некоторых рибосомных белков. Специфические рибосомные белки *M. jannaschii* имеют гомологов у эукариот, но не у эубактерий. Большинство распознанных факторов трансляции у этой архебактерии также оказалось эукариотического

типа. То же, хотя и в меньшей степени, относится к аминоксил-тРНК-синтетазам.

При сравнительном анализе генов системы транскрипции оказалось, что РНК-полимеразы *M. jannaschii* и эубактерий обнаруживают гомологию среди субъединиц, формирующих минимальный фермент, однако архебактерия обладает малыми дополнительными субъединицами, которые не свойственны эубактериям, а их гомологи имеются у РНК-полимераз эукариот. Лишь два из основных факторов транскрипции *M. jannaschii* гомологичны таковым эукариот, а один или два фактора рассматриваются, как "рудиментарные" формы соответствующих эукариотических факторов. Таким образом, система транскрипции архебактерий сегодня представляется как более простая и, возможно, более примитивная версия соответствующей эукариотической системы.

В геноме *M. jannaschii* найден только один ген, кодирующий ДНК-полимеразу, которая напоминает эукариотическую ДНК-полимеразу ϵ . ДНК-полимераза Pol III, осуществляющая репликацию ДНК у эубактерий, не имеет гомолога у *M. jannaschii*. Высокую гомологию с белками эукариот обнаруживают и другие белки архебактерии: гистоны, белки, контролирующие деление клетки, протеасомы, факторы элонгации трансляции, белки систем репарации и транспорта. Для *M. jannaschii*, как и для эубактерий, характерна организация генов в виде оперонов. Однако в первом случае опероны встречаются редко и почти всегда объединяют гены субъединиц белковых комплексов, например РНК-полимеразы, рибосом или метил-коэнзим М-редуктазы. В то же время довольно редки опероны, содержащие гены, объединенные по принципу контроля последовательных метаболических реакций. У *M. jannaschii* такие гены могут быть случайным образом распределены по геному.

Итак, несмотря на то что архебактерии образуют особое царство и по ряду своих генетических свойств приближаются к эукариотам, размер их генома и набор основных генов остаются типичными для свободно живущих бактерий.

1.2.4. Минимальный размер генома одноклеточных организмов

Определение минимального размера генома, обеспечивающего все необходимые функции, которые позволяют одноклеточному организму

существовать в определенных экологических условиях, не является праздным вопросом. Решение этой проблемы необходимо для понимания происхождения жизни на Земле, а также путей и механизмов совместного эволюционирования генов, объединенных в конкретные геномы, а следовательно, и механизмов возникновения геномов как таковых. Данная проблема была впервые сформулирована Дж. Холдейном в 1920-е годы и с тех пор неоднократно исследовалась. Недавнее (1995 г.) определение полных первичных структур ДНК геномов двух паразитических микроорганизмов (*Mycoplasma genitalium* и *Haemophilus influenzae*) дало возможность использовать новый подход для изучения данной проблемы. А.Р. Мушегианом и Е.В. Куниным проведен детальный сравнительный анализ полного набора генов этих микроорганизмов, который позволил составить перечень генов, абсолютно необходимых для существования свободно живущих клеток.

Считается, что геномы *M. genitalium* и *H. influenzae* произошли путем последовательного уменьшения размера генома соответственно грамположительных и грамотрицательных бактерий-предшественников с более крупными геномами после отделения их предков от общего предшественника не менее 1,5 миллиардов лет назад. Предполагается, что общие гомологичные гены, сохранившиеся у этих микроорганизмов на протяжении столь длительного периода их существования, являются жизненно важными и составляют основу минимального набора генов, необходимых для автономного существования паразитических клеток. На основе анализа последовательностей нуклеотидов предсказано, что геном *M. genitalium*, длина которого составляет ~580 т.п.о., кодирует 469 белков, тогда как геном *H. influenzae* (~1830 т.п.о.) кодирует 1703 белка. Оказалось, что геном *M. genitalium* содержит в своем составе 240 генов, имеющих функциональные гомологи в геноме *H. influenzae*. При теоретической разработке идеального абстрактного генома к набору было добавлено 22 гена, необходимых для осуществления жизненно важных метаболических процессов, которые у этих двух микроорганизмов контролировались негомологичными генами. Одновременно из набора были удалены 6 генов, избыточных с точки зрения выполняемых ими функций, которые обеспечивают специфическое взаимодействие микроорганизмов с хозяевами. Оставшиеся 256 генов, по мнению авторов теоретической разработки, полностью перекрывают

потребности абстрактного паразитического микроорганизма. Предлагаемый гипотетический минимальный набор генов, кодирующих 256 белков, должен включать следующие жизненно важные генетические системы микроорганизмов: **почти полный набор генов системы трансляции; почти полный набор генов системы репликации; гены рудиментарной системы репарации и рекомбинации; гены аппарата транскрипции, в котором отсутствуют почти полностью гены регуляции транскрипции; большой набор генов, кодирующих белки, гомологичные шаперонам; гены, контролирующие анаэробный метаболизм, включая гены гликолиза и фосфорилирования субстратов; гены биосинтеза липидов; восемь генов, кодирующих ферменты, которые используют сложные кофакторы; гены системы транспорта белков; ограниченный набор генов, обеспечивающий транспорт метаболитов; полный набор генов утилизации нуклеотидов *de novo* и гены их биосинтеза; гены биосинтеза аминокислот не включены (поскольку предполагается паразитический образ жизни).**

Ранее были использованы еще два подхода для определения минимального размера генома, необходимого для автономного существования микроорганизмов. В одном из них путем введения случайных мутаций определялось число генетических локусов у *Bacillus subtilis*, несущественных для ее выживания. На основании результатов этих исследований сделан вывод о том, что средний размер минимального генома составляет 318 т.п.о., а его максимальный размер приближается к 562 т.п.о. Полученные значения согласуются с величинами, характерными для *M. genitalium*. При другом подходе изучались изменения размера генома при переходе от свободно живущих клеток к облигатным внутриклеточным паразитам и органеллам эукариот. При этом риккетсии рассматривались как эволюционные предшественники митохондрий. Работа еще не завершена, поскольку полной первичной структуры генома риккетсий пока не получено.

Таким образом, исследования, проведенные на геномах *M. genitalium* и *H. influenzae*, дают в настоящее время наиболее точную оценку минимального набора генов (~ 250), необходимых для существования микроорганизмов. Эти результаты будут корректироваться по мере накопления экспериментальных данных о структуре других геномов, что позволит в каждом конкретном случае определять набор именно тех генов, которые делают любой организм

уникальным и неповторимым.

1.3. Геном эукариот

Как уже упоминалось выше, в отличие от прокариот основная часть генома эукариот находится в специальном клеточном компартменте (органелле), получившем название **ядра**, а значительно меньшая часть – в митохондриях, хлоропластах и других пластидах. Так же, как и у прокариот, информационной макромолекулой генома эукариот является ДНК, которая неравномерно распределена по нескольким хромосомам в виде комплексов с многочисленными белками. Эти ДНК-белковые комплексы эукариот получили название **хроматина**. На протяжении клеточного цикла хроматин претерпевает высокоупорядоченные структурные преобразования в виде последовательных конденсаций–деконденсаций. В соматических клетках при максимальной конденсации в метафазе митоза эти преобразования сопровождаются формированием видимых в микроскопе **метафазных** хромосом. Как морфология метафазных хромосом, так и их число являются уникальными характеристиками вида. Совокупность внешних признаков хромосомного набора эукариот получила название **кариотипа**. Эти признаки широко используются в биологической систематике.

Геном эукариот существенно отличается от генома прокариот по ряду признаков, среди которых необходимо отметить его избыточность. Содержание ДНК у эукариот в расчете на одну клетку в среднем на два–три порядка выше, чем у прокариот, и у разных видов животных изменяется от 168 пг (амфибии) до 1 пг (некоторые виды рыб). У человека имеется ~ 6 пг ДНК на диплоидный геном, суммарная длина которой приближается к $6 \cdot 10^9$ п.о. (см. табл. I.1).

Повышенное содержание ДНК в геноме эукариот нельзя объяснить одним лишь увеличением потребности этих организмов в дополнительной генетической информации в связи с усложнением организации, поскольку большая часть их геномной ДНК, как правило, представлена некодирующими последовательностями нуклеотидов. Размер генома организмов, находящихся на более низких ступенях эволюционного развития, зачастую превышает размеры геномов более высокоорганизованных животных и растений. В настоящее время известно, что большая часть ДНК генома эукариот не

кодирует РНК и белки, и ее генетические функции не вполне понятны. Особенности первичной структуры ДНК эукариот позволяют разделить ее на многочисленные семейства и классы, основные из которых кратко рассмотрены ниже.

1.3.1. Последовательности нуклеотидов эукариотического генома

Геном эукариот составляют уникальные и повторяющиеся последовательности нуклеотидов. Содержание уникальных последовательностей в геноме, определенное на основании кинетики реассоциации фрагментированной ДНК, варьирует у разных организмов, и их доля составляет 15–98% от всей ДНК. Несмотря на то, что во фракцию уникальных последовательностей попадают многие структурные гены, большая часть уникальных последовательностей является некодирующей и обычно не заключает в себе генетической информации в общепринятом значении этого термина: не кодирует функционально значимые полипептидные цепи или РНК. Хорошо известным примером таких уникальных последовательностей являются интроны, общий размер которых, как правило, на порядок и более превышает суммарный размер экзонов содержащих их генов.

Эволюционное возникновение мозаичной (интрон–экзонной) структуры генов эукариот, так же как и консервативный характер наследования размеров и взаимного расположения интронов в генах, не находит в настоящее время исчерпывающего объяснения из-за кажущегося отсутствия фактора давления естественного отбора на последовательности нуклеотидов без четких биологических функций. Наибольшее распространение получила концепция **В. Гилберта (1977 г.), согласно которой появление интронов, по-видимому, совпавшее по времени с эволюционным возникновением многоклеточных организмов, обеспечило возможность обмена экзонами между неродственными генами (exon shuffling)**. Такой обмен должен сопровождаться образованием новых белков мозаичного строения, составленных из готовых полипептидных функционально значимых модулей (доменов), ранее принадлежавших другим белкам. Следствием этого, по мнению сторонников данной концепции, было резкое ускорение образования белков и ферментов с новыми функциями, а также глубокие эволюционные преобразования самих организмов, реализующих такие молекулярные

механизмы. Эта точка зрения получила название "**гипотезы позднего возникновения интронов**" (intron late). В соответствии с другой гипотезой Дж.Е. Дарнелла и В.Ф. Дулиттла (1978 г.) современные интроны представляют собой "**эволюционные реликты**". Когда-то интроны были частью гигантских генов.

Не менее загадочным с эволюционной точки зрения остается и феномен появления в геноме многоклеточных организмов большого количества некодирующих повторяющихся последовательностей. Такие повторы представлены в гаплоидном геноме эукариот множественными копиями. В современной классификации повторов принято различать **часто повторяющиеся последовательности, число которых превышает 10^5 на гаплоидный геном, и умеренно повторяющиеся, представленные $10-10^4$ копиями**. Хорошо изученным представителем первых является **сателлитная ДНК**, которая состоит из коротких tandemных повторов длиной 1–20 п.о., организованных в длинные блоки. Одними из первых среди повторяющихся последовательностей ДНК эукариот были открыты сателлитные ДНК тимуса телят. Свое название они получили на основании того, что при анализе суммарной эукариотической ДНК центрифугированием в градиенте плотности хлористого цезия они сопровождали основной пик оптической плотности в виде плеча (спутника, сателлита). Именно гомогенный нуклеотидный состав фракции сателлитных ДНК, определяемый наличием в ней многочисленных коротких повторов, изменял ее плавучую плотность, что легко обнаруживалось при центрифугировании. В своем классическом определении сателлитных ДНК Р.Д. Бриттен и соавт. (1974 г.) отмечали, что **сателлиты – это минорный компонент ДНК, отделяющийся от основной ДНК при равновесном ультрацентрифугировании в градиенте плотности CsCl**. Для сателлитов характерен ряд свойств, среди которых наиболее важны: а) быстрая и точная реассоциация в процессе ренатурации ДНК; б) множество копий; в) простая первичная структура; г) гомогенный состав (протяженные кластеры одних и тех же повторяющихся блоков последовательны); д) пурип–пиримидиновая асимметрия в распределении нуклеотидов по цепям ДНК; е) концентрирование в прицентромерном гетерохроматине; ж) ограниченная репликация (недорепликация) при политенизации хромосом; з) нахождение в составе хромосом в виде tandemно (друг за другом) расположенных кластеров.

Содержание сателлитной ДНК в геноме эукариот может достигать 5–50% от суммарного количества ДНК. *Микро-* (от 1 до 4 п.о. в основном повторяющемся блоке) и *минисателлитные* (с бóльшим числом п.о. в индивидуальном повторе) ДНК характеризуются высокой вариабельностью по числу копий в геномах организмов даже одного вида и в ряде случаев обладают генетической нестабильностью как в норме, так и при некоторых патологических состояниях организмов. Благодаря этому свойству мини- и микросателлиты часто называют тандемными повторами с изменяющимся числом копий **VNTR (variable number of tandem repeats)**.

Другой тип повторов – **диспергированные повторяющиеся последовательности ДНК**, не организованные в крупные блоки, а рассеянные по геному. Повторы этого типа, иначе называемые умеренно повторяющимися последовательностями (*medium reiterated frequency repeats – MERs*), разделяют на два обширных класса: **SINE** (*short interspersed elements*) – короткие и **LINE** (*long interspersed elements*) – длинные диспергированные элементы. Длина SINE-элементов составляет 90–400 п.о., тогда как длина LINE-последовательностей может достигать 7 т.п.о. Хорошо изученными повторами класса SINE в геноме человека и некоторых приматов являются так называемые **Alu-повторы**, длина повторяющейся единицы которых составляет ~300 п.о. Alu-повторы представлены в геноме человека ~10⁶ копиями и в среднем встречаются через каждые 4 т.п.о., составляя ~5% от суммарного количества ДНК. Аналогичные в структурном отношении повторы, названные B1, обнаружены в геноме мышей и под другими названиями описаны у многих млекопитающих.

Хотя **LINE-последовательности** *заклюают в себе гены обратных транскриптаз*, что является признаком ретротранспозонов (мобильных генетических элементов животных, обладающих структурным сходством с геномом ретровирусов), для них характерно отсутствие последовательностей длинных концевых повторов (*long terminal repeats – LTR*), типичных для ретротранспозонов (подробнее о геноме ретровирусов см. раздел 7.2.7). В качестве примера LINE-последовательности можно упомянуть LINE-1-повтор, широко распространенный в геноме животных. LINE-1-элемент мышей содержит две открытые рамки считывания ORF-1 и ORF-2, вторая из которых

кодирует белок, гомологичный обратной транскриптазе. ORF фланкированы короткими нетранслируемыми последовательностями, а сами LINE-1 – короткими прямыми повторами (SDR). 5'-Концевые последовательности повтора функционируют в качестве промоторов транскрипции. Этот участок LINE-1 грызунов (но не человека) построен из коротких tandemных повторов двух типов А и F, называемых мономерами. Длина мономеров у крыс составляет 600 п.о. При этом А- (но не F) мономеры обладают активностью промоторов.

Так же как и сателлитные ДНК, SINE- и LINE-повторы характеризуются генетической нестабильностью. Их общими чертами являются транскрибируемость и способность к транспозициям. Последовательности РНК, транскрибированные с умеренных повторов, обнаруживают среди гетерогенных ядерных РНК, где их доля достигает 20–30%. Имеются экспериментальные свидетельства того, что новые копии повторяющихся элементов обоих типов возникают в геноме в результате функционирования механизма, названного **ретротранспозицией**, или **ретропозицией**. При участии подобного механизма под действием обратной транскриптазы сначала образуется кДНК на матрице РНК-транскрипта соответствующего повтора, которая далее интегрируется в новый локус генома, как это имеет место у ретровирусов. Такой механизм дает возможность локально изменять число копий определенных последовательностей нуклеотидов в эукариотическом геноме. Тем не менее, большая часть LINE-последовательностей неспособна к транспозициям, и их ORF, по-видимому, могут быть отнесены к *псевдогенам* – неэкспрессирующимся последовательностям, гомологичным последовательностям истинных генов. Помимо вышеупомянутых повторяющихся последовательностей геном человека содержит более 100 000 копий MaLR-повторов длиной в 2–3 т.п.о., содержащих LTR, и несколько тысяч последовательностей генома ретровирусов.

Несмотря на широкую распространенность повторяющихся и уникальных некодирующих последовательностей в геноме эукариот и их очевидную активность во время жизненного цикла организмов, биологическое значение этих и других некодирующих элементов генома остается непонятным. Вызывает сомнение правильность активно обсуждаемой в литературе гипотезы об "эгоистичности" избыточной геномной ДНК, в соответствии с которой вся

избыточная ДНК является геномным паразитом и распространяется в геноме в результате транспозиций точных копий немногочисленных исходных последовательностей. Действительно, слишком велики были бы энергетические затраты на биосинтез предшественников ДНК и самой ДНК в клетках, в которых содержание "паразитической" ДНК в геноме на 2–3 порядка превышает количество функционально значимой ДНК, заключающей в себе последовательности нуклеотидов генов. Клетки с геномом, "зараженным" эгоистической ДНК, не смогли бы выдерживать конкуренции с клетками, не содержащими "паразита", из-за значительного возрастания энергетических затрат на редупликацию генома. Кроме того, концепция эгоистической ДНК, в соответствии с которой предполагается отсутствие эволюционного давления отбора на "паразитические" последовательности нуклеотидов, не объясняет высокую консервативность мест локализации и размеров интронов в гомологичных генах филогенетически близких организмов, а также не указывает на механизм, поддерживающий число копий повторов на относительно постоянном уровне в ряду поколений организмов. Такого рода концепции не могут ответить на вопрос: где же предел, до которого может самопроизвольно увеличиваться размер генома клетки-хозяина в филогенезе?

Функциональную значимость избыточной ДНК лишь частично объясняют концепции, приписывающие ей структурную роль в пространственной организации генома и участие в конъюгации гомологичных хромосом в мейозе или репликации теломерных участков хромосом. Таким образом, основные положения знаменитого парадокса С, указывающие на необъяснимое присутствие в геноме эукариотических организмов большого количества избыточной ДНК, по-прежнему остаются загадочными и парадоксальными. Попытка нового объяснения основной функциональной роли "избыточной" ДНК в геноме эукариот сделана в разделе 5.3.

1.3.2. Хроматин

Хроматином называют сложную смесь веществ, из которых построены хромосомы эукариот. Основными компонентами хроматина являются ДНК, гистоны и негистоновые белки, образующие высокоупорядоченные в пространстве структуры. Соотношение ДНК и белка в хроматине составляет ~1:1, а основная масса белка хроматина представлена гистонами. Гистоны

образуют семейство высококонсервативных основных белков, которые разделяются на пять больших классов, названных **H1**, **H2A**, **H2B**, **H3** и **H4**. Размер полипептидных цепей гистонов лежит в пределах **~220 (H1)** и **102 (H4)** аминокислотных остатков. Гистон H1 сильно обогащен остатками **Lys**, для гистонов H2A и H2B характерно умеренное содержание Lys, полипептидные цепи гистонов H3 и H4 богаты **Arg**. Внутри каждого класса гистонов (за исключением H4) на основании аминокислотных последовательностей различают несколько субтипов этих белков. Такая множественность особенно характерна для гистонов класса H1 млекопитающих. В этом случае различают семь субтипов, названных H1.1–H1.5, H1⁰ и H1t.

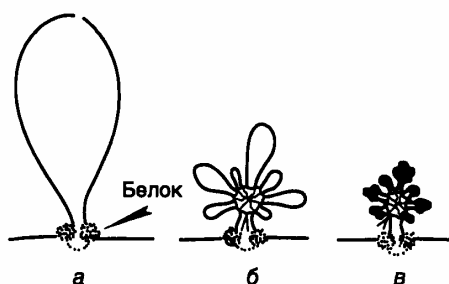


Рис. 1.2. Схематическое изображение петельно-доменного уровня компактизации хроматина

а – фиксация петли хромомера на ядерном матриксе с помощью MAR/SAR-последовательностей и белков; б – "розетки", образованные из петли хромомера; в – конденсация петель "розеток" с участием нуклеосом и нуклеомеров

Важным результатом взаимодействия ДНК с белками в составе хроматина является ее компактизация. Суммарная длина ДНК, заключенной в ядре клеток человека, приближается к 1 м, тогда как средний диаметр ядра составляет ~ 10 мкм. Длина молекулы ДНК, заключенной в одной хромосоме человека, в среднем равняется ~ 4 см. В то же время длина метафазной хромосомы составляет ~ 4 мкм. Следовательно, ДНК метафазных хромосом человека компактизована по длине, по крайней мере, в $\sim 10^4$ раз. Степень компактизации ДНК в интерфазных ядрах значительно ниже и неравномерна в отдельных генетических локусах. С функциональной точки зрения различают **эухроматин** и **гетерохроматин**. Эухроматин характеризуется меньшей по сравнению с гетерохроматином компактизацией ДНК, и в нем главным образом локализуются активно экспрессирующиеся гены. В настоящее время широко

распространено мнение о генетической инертности гетерохроматина. Поскольку его истинные функции сегодня нельзя считать установленными, эта точка зрения по мере накопления знаний о гетерохроматине может измениться. Уже сейчас в нем находят активно экспрессирующиеся гены.

Гетерохроматизация определенных участков хромосом часто сопровождается подавлением транскрипции имеющихся в них генов. В процесс гетерохроматизации могут быть вовлечены протяженные участки хромосом и даже целые хромосомы. В соответствии с этим считается, что регуляция транскрипции генов эукариот в основном происходит на двух уровнях. На первом из них компактизация или декомпактизация ДНК в хроматине может приводить к длительной инактивации или активации протяженных участков хромосом или даже целых хромосом в онтогенезе организма. Более тонкая регуляция транскрипции активированных участков хромосом достигается на втором уровне при участии негистоновых белков, включающих многочисленные факторы транскрипции.

Структурная организация хроматина и хромосом эукариот. Вопрос о структурной организации хроматина в интерфазных ядрах в настоящее время далек от своего разрешения. Это связано, прежде всего, со сложностью и динамичностью его структуры, которая легко меняется даже при незначительных экзогенных воздействиях. Большинство знаний о структуре хроматина было получено *in vitro* на препаратах фрагментированного хроматина, структура которого значительно отличается от таковой в нативных ядрах. В соответствии с распространенной точкой зрения различают три уровня структурной организации хроматина у эукариот: 1) **нуклеосомная фибрилла**; 2) **соленоид, или нуклеомер**; 3) **петельно-доменная структура, включающая хромомеры**.

Нуклеосомные фибриллы. В определенных условиях (при низкой ионной силе и в присутствии двухвалентных ионов металлов) в изолированном хроматине удастся наблюдать регулярные структуры в виде протяженных фибрилл диаметром 10 нм, состоящих из нуклеосом. Эти фибриллярные структуры, в которых нуклеосомы расположены как бусы на нитке, рассматриваются в качестве низшего уровня упаковки ДНК эукариот в хроматине. Нуклеосомы, входящие в состав фибрилл, расположены более или менее равномерно вдоль молекулы ДНК на расстоянии 10–20 нм друг от друга.

В состав нуклеосом входят четыре пары молекул гистонов: H2a, H2b, H3 и H4, а также одна молекула гистона H1. Данные по структуре нуклеосом в основном получены с использованием трех методов: рентгеноструктурного анализа низкого и высокого разрешения кристаллов нуклеосом, межмолекулярных сшивок белок–ДНК и расщепления ДНК в составе нуклеосом с помощью нуклеаз или радикалов гидроксила. На основании таких данных А. Клугом была построена модель нуклеосомы, в соответствии с которой ДНК (146 п.о.) в *В-форме* (правозакрученная спираль с шагом 10 п.о.) намотана на гистоновый октамер, в центральной части которого расположены гистоны H3 и H4, а на периферии – H2a и H2b. Диаметр такого нуклеосомного диска составляет 11 нм, а его толщина – 5,5 нм. Структура, состоящая из гистонического октамера и намотанной на него ДНК, получила название *нуклеосомной кóровой частицы*. Кóровые частицы отделены друг от друга сегментами *линкерной ДНК*. Общая длина участка ДНК, включенного в нуклеосому животных, составляет 200 (± 15) п.о.

Полипептидные цепи гистонов содержат структурные домены нескольких типов. Центральный глобулярный домен и гибкие выступающие N- и C-концевые участки, обогащенные основными аминокислотами, получили название *плеч* (arm). C-концевые домены полипептидных цепей, участвующие в гистон–гистоновых взаимодействиях внутри кóровой частицы, находятся преимущественно в виде α -спирали с протяженным центральным спиральным участком, вдоль которого с двух сторон уложено по одной более короткой спирали. Все известные места обратимых посттрансляционных модификаций гистонов, происходящих на протяжении клеточного цикла или во время дифференцировки клеток, локализованы в гибких основных доменах их полипептидных цепей (табл. I.2). При этом N-концевые плечи гистонов H3 и H4 являются самыми консервативными участками молекул, а гистоны в целом – одними из наиболее эволюционно консервативных белков. С помощью генетических исследований дрожжей *S. cerevisiae* было установлено, что небольшие делеции и точковые мутации в N-концевых частях генов гистонов сопровождаются глубокими и разнообразными изменениями фенотипа дрожжевых клеток. Это указывает на чрезвычайную важность целостности молекул гистонов в обеспечении правильного функционирования

эукариотических генов.

В растворе гистоны H3 и H4 могут существовать в виде стабильных тетрамеров $(H3)_2(H4)_2$, а гистоны H2A и H2B – в виде стабильных димеров. Постепенное повышение ионной силы в растворах, содержащих нативный хроматин, приводит к освобождению сначала димеров H2A/H2B, а затем тетрамеров H3/H4.

Дальнейшее уточнение тонкой структуры нуклеосом в кристаллах было проведено недавно в работе К. Люгера с соавт. (1997 г.) с помощью рентгеноструктурного анализа высокого разрешения. Было установлено, что выпуклая поверхность каждого гистонового гетеродимера в составе октамера огибается сегментами ДНК длиной 27–28 п.о., расположенными по отношению друг к другу под углом 140° , которые разделены линкерными участками длиной в 4 п.о.

В соответствии с современными данными пространственная структура ДНК в составе коровых частиц несколько отличается от В-формы: двойная спираль ДНК перекручена на 0,25–0,35 п.о./виток двойной спирали, что приводит к образованию шага спирали, равному 10,2 п.о./виток (у В-формы в растворе – 10,5 п.о./виток). Стабильность комплекса гистонов в составе коровой частицы определяется взаимодействием их глобулярных частей, поэтому удаление гибких плеч в условиях мягкого протеолиза не сопровождается разрушением комплекса. N-концевые плечи гистонов, по-видимому, обеспечивают их взаимодействие со специфическими участками ДНК. Так, N-концевые домены гистона H3 контактируют с участками ДНК на входе в коровую частицу и выходе из нее, тогда как соответствующий домен гистона H4 связывается с внутренней частью ДНК нуклеосомы.

Упомянутые выше исследования структуры нуклеосом высокого разрешения показывают, что центральная часть сегмента ДНК длиной в 121 п.о. в составе нуклеосомы образует дополнительные контакты с гистоном H3. При этом N-концевые части полипептидных цепей гистонов H3 и H2B проходят через каналы, образуемые малыми бороздками соседних супервитков ДНК нуклеосомы, а N-концевая часть гистона H2A контактирует с малой бороздкой внешней части супервитка ДНК. В совокупности данные высокого разрешения показывают, что ДНК в составе коровых частиц нуклеосом огибает гистоновые октамеры неравномерно. Кривизна нарушается в местах взаимодействия ДНК с

поверхностью гистонов, и такие изломы наиболее заметны на расстоянии 10–15 и 40 п.о. от центра супервитка ДНК.

Таблица I.2

Свойства гистонов животных

Гистон	Размер полипептида (число аминокислот)			Локализация и типы посттрансляционных модификаций
	Вся молекула	N-концевое плечо	C-концевое плечо	
H3	135	41	25	Ac-K9, P-S10, Ac-K-14, Ac-K18, Ac-K23, Ac-K27
H4	102	32	–	P-S1, Ac-K5, Ac-K8, Ac-K12, Ac-K16, Ac-K20
H2A	129	24	16	P-S1, Ac-K5, Ac-K9
H2B	125	30	23	Ac-K12, Ac-K15, Ac-K20, Ac-K24, P-S32, P-S36
H1	213	36	92	S-фаза: P-S145, P-S173, P-S180 Митоз: P-S16, P-T17, P-T136, P-S145, P-S173, P-S180

Примечание. Ac – ацетилирование, P – фосфорилирование; K, R, T и S – остатки Lys, Arg, Thr и Ser соответственно. Цифры обозначают положение соответствующих аминокислотных остатков в полипептидных цепях.

Структура фибриллы хроматина типа "бусин на нитке" наблюдается *in vitro* при очень низкой ионной силе и в отсутствие двухвалентных катионов или же без гистона H1 даже при физиологических концентрациях солей. При повышении ионной силы в присутствии гистона H1, который выполняет функции конденсирующего агента, в хроматине происходит структурный переход, сопровождающийся формированием равномерной фибриллы диаметром 30 нм, в которой нуклеосомы располагаются вплотную друг к другу. Предполагается, что именно такая структура хроматина на этом уровне конденсации преобладает *in vivo*.

В ядрах ДНК в основном входит в состав фибрилл диаметром 25~30 нм,

которые образуются и в растворах с физиологическими концентрациями солей (~100 мМ NaCl), содержащих изолированный хроматин. Такие фибриллы в форме соленоида неравномерны по своей длине. С помощью электронного микроскопа в них обнаруживаются гранулы диаметром 30 нм, между которыми находятся более тонкие участки. Ю.С. Ченцов, впервые обнаруживший эти гранулы, назвал их *нуклеомерами*, позднее они были описаны в литературе под названием "*супербидов*" (super-beads). Согласно модели А. Клуга, на один шаг (10 нм) соленоида приходится ~6 нуклеосом, что приводит к уменьшению в ~6 раз длины фибриллы диаметром в 10 нм, а исходной свободной ДНК – в 40 раз. Таким образом, в соответствии с этой упрощенной картиной цепь нуклеосом на втором уровне упаковки хроматина свернута в симметрично построенный соленоид, содержащий нуклеомеры. Симметрия соленоида нарушается при взаимодействии с ним негистоновых белков.

Энергия, необходимая для образования соленоида, частично черпается из энергии взаимодействия нуклеосом друг с другом. Однако основной вклад в этот процесс вносит гистон Н1, связывание молекул которого с линкерной ДНК за пределами *кóровых* частиц обеспечивает образование компактных фибрилл диаметром 30 нм. Дифференцирующиеся клетки высших эукариот обладают способностью к синтезу целого семейства гистонов Н1, например Н1А, Н1В, Н1^о, которые имеют структурное сходство с гистоном Н5, обнаруженным в дифференцирующихся предшественниках эритроцитов. Полипептидные цепи гистонов семейства Н1, как и молекулы гистонов *кóровых* частиц, содержат центральный глобулярный домен и два гибких плеча. Глобулярный домен определяет локализацию гистона Н1 в нуклеосомах, где он взаимодействует с ДНК в месте ее входа в *кóровую* частицу и выхода из нее, стабилизируя два отрицательных супервитка ДНК, обернутой вокруг гистонов *о*ктамера. Расщепление хроматина нуклеазой микрококков приводит к освобождению гистона Н1, так как при этом происходит отщепление сегментов ДНК, с которыми он взаимодействует.

Следует подчеркнуть, что точная структура соленоида в настоящее время неизвестна. Данные, недавно полученные методом рассеивания нейтронов, указывают на то, что гистон Н1 локализован внутри хроматиновой фибриллы и на ее поперечном срезе помещается шесть нуклеосом, что хорошо

соответствует модели Клуга. Однако при изучении хроматина с помощью сканирующей микроскопии было установлено, что при низкой ионной силе соленоид существует в виде нерегулярной спирали.

Петельно-доменный уровень компактизации хроматина. В интерфазных ядрах эукариот нити хроматина, в которых ДНК упакована в форме соленоида, содержащего нуклеомеры, организованы в виде топологически независимых петель, длина которых в среднем составляет 50–100 т.п.о. Такой способ пространственной укладки нитей хроматина рассматривается как следующий уровень конденсации хроматина (и ДНК) у эукариот, а сами петли получили название *хромомеров*. С помощью электронного микроскопа обнаружено, что нити хроматина в хромомерах имеют дополнительную специфическую укладку в виде розеток, собранных у основания, от которого отходят малые петли длиной ~5 т.п.о. (см. рис. 1.2). Образование хромомеров становится возможным благодаря наличию у их оснований определенных последовательностей нуклеотидов, которые специфически взаимодействуют с *ядерным матриксом*, называемым иначе ядерным *скэффолдом* (скелетом) – сетчатообразной структурой внутри интерфазных ядер, образованной белками и РНК. Эти участки хромосомной ДНК, взаимодействующие с ядерным матриксом, в литературе известны под сокращенными названиями MAR (Matrix Associated Region) или SAR (Scaffold Associated Region) и часто обозначаются как MAR/SAR-последовательности.

Были предприняты попытки связать локализацию MAR/SAR-последовательностей на ДНК с функциональной активностью хроматина. Исследование петельной структуры хроматина в ядрах дрозофилы показало, что эти последовательности часто располагаются в окрестностях генов, транскрибируемых РНК-полимеразой II, и фланкируют участки хроматина, содержащие одну или несколько транскрипционных единиц. Выяснилось, что активно транскрибируемые гены организованы в петли небольшого размера (~10 т.п.о.), тогда как "молчащие" гены находятся преимущественно в составе более крупных петель, которые содержат несколько транскрипционных единиц и, в свою очередь, образуют дополнительные петли в виде розеток (см. рис. 1.2). Как правило, MAR/SAR-последовательности фланкируют гены, однако в ряде случаев их обнаруживают и внутри генов, но в составе интронов. В настоящее время выделено и охарактеризовано несколько десятков MAR/SAR-

последовательностей из генома дрозофилы и других организмов. У дрозофилы они представляют собой АТ-богатые последовательности нуклеотидов длиной в 200–350 п.о., в составе которых часто встречаются участки поли(dA)-поли(dT). Обнаружены два типа блоков из АТ-богатых последовательностей: А-блок – ААТАААТ/СААА и Т-блок – ТТА/ТТТ/АТТТ/СААА, характерные для всех MAR/SAR-последовательностей не только дрозофилы, но и других организмов. Отличающиеся от них блоки похожих последовательностей нуклеотидов – АТАТТТ и ААТАТТТТ, первоначально найденные в MAR/SAR-последовательностях мышей, оказались широко представленными и в аналогичных последовательностях многих эукариот.

Можно выделить следующие обобщенные характеристики MAR/SAR-последовательностей, приведенные в недавнем обзоре М.В. Глазкова (1995 г.):

длина MAR/SAR-последовательностей составляет 300–1000 п.о.; все они содержат многочисленные сайты ДНК-белкового взаимодействия и обогащены АТ-парами;

MAR/SAR-последовательности обладают способностью обратимо связываться с ядерным матриксом метафазных хромосом и локализованы исключительно в некодирующих последовательностях генома, главным образом в нетранскрибируемых участках, реже в интронах;

расстояние между двумя соседними MAR/SAR-последовательностями составляет 3–112 т.п.о., и они фланкируют один или несколько экспрессирующихся генов, для которых характерна повышенная чувствительность к нуклеазам из-за деконденсированного состояния их хроматина;

некоторые MAR/SAR-последовательности выявлены рядом с энхансероподобными регуляторными элементами. Кроме того, эти последовательности могут быть потенциальными сайтами инициации репликации и содержать в своем составе большое число сайтов для разных факторов транскрипции;

MAR/SAR-последовательности иногда соседствуют с последовательностями, способными образовывать Z- или H-формы ДНК, и в таких случаях они определяются как сайты с повышенной чувствительностью к ДНКазе I.

Все это указывает на то, что MAR/SAR-последовательности являются

чрезвычайно важными в функциональном отношении последовательностями геномной ДНК эукариот. Одной из основных функций, которую им приписывают в настоящее время, может быть пространственное разграничение функциональных доменов хроматина в интерфазных ядрах эукариот, необходимое для эффективной и независимой экспрессии генов, находящихся в этих доменах. Подробнее о выполнении MAR/SAR-последовательностями функций пограничных последовательностей (инсуляторов), регулирующих экспрессию эукариотических генов, см. в разделе 3.2.4.

Являясь одними из ключевых *цис*-действующих генетических регуляторных элементов хромосом эукариот, MAR/SAR-последовательности осуществляют глобальный контроль изменений структуры хроматина и связанных с ними модуляций экспрессии генов на уровне транскрипции. В настоящее время считается, что эти последовательности обеспечивают поддержание функциональной структуры хромосом в интерфазных ядрах и вовлечены в процесс их конденсации при формировании метафазных хромосом во время митоза. Современная модель структуры метафазной хромосомы К.М. Харта и У.К. Лэммли (1998 г.) подчеркивает, что SAR-последовательности, пространственно примыкая друг к другу, образуют ось хроматиды, от которой в разные стороны отходят петли хроматина, формирующие тело метафазной хромосомы. Однако еще остается доказать, являются ли MAR/SAR-последовательности местами специфического прикрепления белковых комплексов, включающих, к примеру, белок *конденсин*, которые необходимы для конденсации хроматина.

Негистоновые белки хроматина. Большое влияние на структуру хроматина и функционирование эукариотических генов оказывают различные *негистоновые белки*. В ядрах в наибольшем количестве обнаруживают негистоновые белки, которые относят к так называемой *группе белков с высокой подвижностью* (high mobility group – HMG). Это название отражает их высокую подвижность при электрофорезе. Суммарное содержание HMG-белков в ядрах клеток в ~10 раз меньше, чем гистонов. Эти белки разделяют на три основных подкласса: 14/17, 1/2 и I/Y.

HMG 14/17 представляют собой небольшие белки с молекулярной массой 10–12 кДа, полипептидные цепи которых несут большой локальный электрический заряд при физиологических значениях pH. N-Концевые части их

полипептидных цепей обладают основными свойствами, тогда как С-концевые – кислыми. Кóровые частицы нуклеосом содержат на своей поверхности две молекулы HMG 14/17, которые соединяют между собой цепи ДНК двух соседних витков. HMG-Белки этой группы могут замещать гистон H1 в активно транскрибируемых генах. *In vivo* ~10% коровых частиц нуклеосом могут содержать белки HMG 14/17.

Белки группы HMG 1/2 также являются одними из преобладающих ядерных белков с молекулярной массой 25–30 кДа. Эти белки специфически связываются с одноцепочечными участками ДНК, а также участками ДНК, образующими крестообразные структуры в местах, содержащих палиндромные последовательности. Это указывает на их возможное участие в процессах рекомбинации, репарации и(или) репликации. Было показано, что белки HMG 1/2 способствуют сборке нуклеосом *in vitro*, однако физиологическое значение этого факта остается непонятным, поскольку многие отрицательно заряженные молекулы в экспериментальных условиях также ускоряют образование нуклеосом.

Белки HMG I/Y преимущественно взаимодействуют с АТ-богатыми участками ДНК, что характерно и для гистона H1. Предполагается, что эти белки конкурируют с гистоном H1 *in vivo* за промоторы и области начала репликации ДНК (см. разделы 2.1 и 4.1).

Помимо HMG-белков к негистоновым белкам хроматина относятся многочисленные внутриядерные ферменты и белковые факторы, необходимые для работы генетического аппарата клетки. Многим из этих белков, обеспечивающим транскрипцию, репарацию и репликацию, посвящены целые разделы этой книги. Среди негистоновых белков особое место занимают ДНК-топоизомеразы, механизмы функционирования которых целесообразно рассмотреть отдельно.

1.3.3. Роль ДНК-топоизомераз в обеспечении структуры и функционирования хроматина

Топоизомеразы контролируют в клетках уровень суперскрученности ДНК, который может изменяться в процессе ее репликации, транскрипции, гомологичной рекомбинации, а также во время перестроек хроматина. Все эти

ферменты релаксируют суперскрученные молекулы ДНК, снимая их внутреннее напряжение путем внесения одно- или двухцепочечных разрывов с последующим их восстановлением (лигированием). По механизму действия различают ДНК-топоизомеразы типа I и II. ДНК-топоизомеразы I, которые являются мономерными белками,

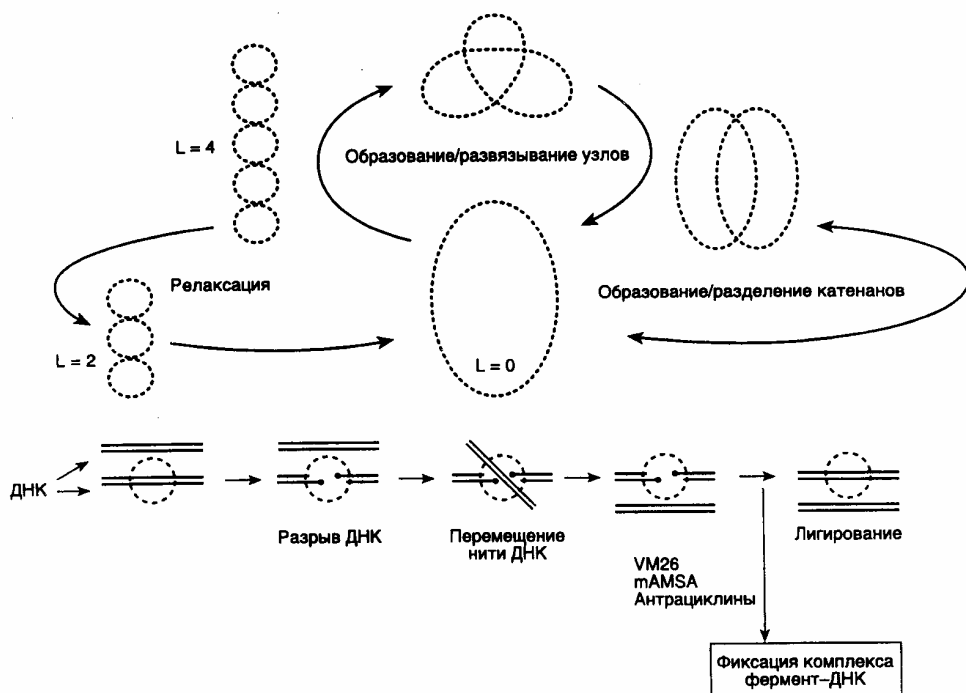


Рис. 1.3. Основные этапы каталитического цикла топоизомеразы II

Внизу изображен механизм переноса нити ДНК через двухцепочечный разрыв. Точками отмечены места ковалентного присоединения фермента к 5'-концам ДНК в двухцепочечных разрывах. VM26, mAMSA, антрациклины – ингибиторы топоизомеразы II, фиксирующие ковалентный комплекс фермент-ДНК и предотвращающие лигирование двухцепочечного разрыва; L – число зацеплений ДНК в суперскрученной молекуле.

релаксируют ДНК без затраты энергии путем внесения одноцепочечных разрывов. В отличие от этого, ДНК-топоизомеразы II функционируют в виде димеров, осуществляя АТФ-зависимое расщепление обеих цепей ДНК с последующим переносом цепей через разрыв и его лигированием (рис. 1.3).

Для внесения одноцепочечного разрыва в ДНК все ДНК-топоизомеразы используют остаток Туг, который осуществляет нуклеофильную атаку фосфатной группы ДНК с образованием фосфотирина. В результате

ферменты оказываются ковалентно связанными с 5'- или 3'-концами ДНК в одноцепочечном разрыве. Образование такой ковалентной связи исключает необходимость затраты энергии при восстановлении фосфодиэфирной связи в одноцепочечном разрыве на заключительных стадиях реакции. У ДНК-топоизомераз типа I имеется один каталитический остаток Туг на молекулу мономерного белка, тогда как димеры ДНК-топоизомераз II содержат по одному каталитическому остатку на каждую субъединицу, что обеспечивает создание ступенчатого двухцепочечного разрыва в молекуле релаксируемой ДНК.

Обнаружены, по крайней мере, два подтипа ДНК-топоизомераз I – IA и IB, которые, будучи неродственными ферментами как по первичной, так и по пространственной структурам, выполняют аналогичные функции с помощью различных механизмов. До недавнего времени ДНК-топоизомеразы IA считали исключительно прокариотическими ферментами, однако они были найдены и в клетках эукариот, включая клетки человека, и названы ДНК-топоизомеразами III.

ДНК-топоизомераза IA релаксирует ДНК, содержащие только отрицательные супервитки, работает в присутствии ионов Mg^{2+} и ковалентно соединяется с 5'-концами ДНК в образующихся временных одноцепочечных разрывах. Это сближает ДНК-топоизомеразы IA и II между собой, что было подтверждено также структурными исследованиями. В отличие от этого, ДНК-топоизомеразы IB способны релаксировать ДНК как с положительными, так и отрицательными супервитками, не требуют для своего функционирования ионов металлов и взаимодействует ковалентно с 3'-концами ДНК. ДНК-топоизомеразы IB найдены исключительно в клетках эукариот (за единственным исключением вируса вакцины).

В клетках человека ДНК-топоизомераза IB/III специфически ингибируется *камптотецином* (camptothecin), который в настоящее время рассматривается в качестве перспективного противоопухолевого препарата. Это соединение взаимодействует преимущественно с ковалентным комплексом топоизомеразы I–ДНК, что подавляет реакцию восстановления фосфодиэфирной связи и освобождение фермента из комплекса. В результате происходит быстрое накопление двухцепочечных разрывов ДНК и вступление клеток в апоптоз.

ДНК-топоизомераза II является жизненно важным ферментом любого

эукариотического организма. Кроме релаксации суперскрученных молекул ДНК она может осуществлять образование или развязывание узлов, а также образование или разделение *катенанов* (кольцевых замкнутых ДНК, сцепленных друг с другом). Реакции развязывания узлов и разделения катенанов являются прерогативой именно ДНК-топоизомеразы II и не выполняются ДНК-топоизомеразами I.

У дрожжей ДНК-топоизомераза II требуется для разделения катенанов сестринских хроматид хромосом в анафазе митоза и абсолютно необходима для сегрегации хромосом в мейозе, а также конденсации хроматина в процессе формирования метафазных хромосом. Выявлена важная роль ДНК-топоизомеразы II и в формировании высших уровней структуры хроматина, а именно участие фермента в образовании петель хроматина во время конденсации хромосом.

ДНК-топоизомераза II локализована в ядре и в больших количествах ассоциируется с ДНК как в интерфазных, так и метафазных ядрах. С помощью специфических антител показано, что молекулы фермента располагаются преимущественно вдоль центральной продольной оси обоих плеч хромосом у многих организмов. Такое аксиальное распределение ДНК-топоизомеразы II в хромосомах наблюдали даже после удаления из них большей части гистонов в результате многократных солевых экстракций. Специфическая локализация этого фермента в хромосомах очень показательна в свете обсуждавшихся выше петельно-доменных особенностей организации хроматина в ядрах. Создается впечатление, что ДНК-топоизомераза II находится в виде гомодимера в основании петель, взаимодействуя с MAR/SAR-последовательностями ДНК хроматина. Хотя топоизомераза II не обнаруживает строгой специфичности в отношении расщепляемых последовательностей нуклеотидов, на выбор сайтов большое влияние оказывают структурные компоненты хроматина. Показано, что *in vivo* существуют два класса сайтов, по которым происходит расщепление ДНК этим ферментом: одни из них локализованы в активно транскрибируемых участках хроматина, гиперчувствительных к действию нуклеаз, а другие – непосредственно в MAR/SAR-последовательностях. Ассоциация ДНК-топоизомеразы II с активно транскрибируемыми участками хроматина указывает на ее возможную важную роль в регуляции экспрессии генов, что и было продемонстрировано в недавних

экспериментах. Таким образом, ДНК-топоизомераза II является одним из ключевых ферментов, необходимых для разрешения сложных топологических проблем, возникающих при изменении структуры хроматина в процессах репликации ДНК, транскрипции генов и сегрегации хромосом в митозе и мейозе. Определенные изоформы (α или β) этого фермента, по-видимому, играют важную роль в поддержании динамической структуры хроматина интерфазных и митотических хромосом.

Имеющиеся в настоящее время данные свидетельствуют о высокоупорядоченной организации генома у любого, особенно эукариотического, организма. Первым уровнем такой упорядоченности является консервативное линейное распределение генов и других последовательностей нуклеотидов вдоль молекул ДНК хромосом, которое служит важным таксономическим признаком. Другим не менее жизненно важным свойством генома эукариот и, по-видимому, таксономическим признаком является его упорядоченное распределение в объеме интерфазных ядер. Высокоспецифическое пространственное распределение хроматина эукариот в интерфазных ядрах можно рассматривать в качестве второго уровня его упорядоченности. Находясь в деконденсированном состоянии после завершения митоза, интерфазные хромосомы не перемешиваются внутри интерфазных ядер, но занимают вполне определенные микрокомпарменты. Определенные участки хромосом, маркированные специфическими (в частности MAR/SAR) последовательностями, служат для прикрепления ДНК хромосом к компонентам ядерного матрикса и ядерных мембран. Такие контакты необходимы для эффективной реализации генетической информации в процессе экспрессии генов, эффективной конденсации хроматина и разделения хромосом в митозе и мейозе. Кроме того, пространственно организованное распределение генетического материала в интерфазных ядрах обеспечивает дифференциальную защиту от мутаций отдельных генетических локусов и, по-видимому, может контролировать темп и направление эволюционных изменений как отдельных локусов, так и организмов в целом (подробнее см. раздел 5.3). К сожалению, исследование пространственной организации генома в интерфазных ядрах (*архитектоники ядра*) сопряжено с большими методическими трудностями и сегодня еще только начинается.

ГЛАВА 2. РЕАЛИЗАЦИЯ ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ ПРИ ЭКСПРЕССИИ ГЕНОВ

Конечным результатом экспрессии генов, кодирующих белки или нуклеиновые кислоты, должно быть образование этих полноценных в функциональном отношении макромолекул, сопровождаемое формированием определенного фенотипа организма. В соответствии с основным постулатом молекулярной биологии генетическая информация в процессе ее реализации передается однонаправленно от нуклеиновых кислот к белкам. При этом реализуется следующая обобщенная схема: ДНК $\leftrightarrow$ РНК $\rightarrow$ белок, которая подчеркивает, что в ряде специальных случаев возможна передача генетической информации от РНК к ДНК с использованием механизма *обратной транскрипции*. До сих пор не обнаружена передача генетической информации от белков к нуклеиновым кислотам. На первом этапе экспрессии генов происходит переписывание генетической информации, заключенной в генах, на *матричные* (информационные) РНК (мРНК – messenger RNA, mRNA), которые являются местом промежуточного хранения этой информации при ее реализации. В некоторых случаях уже сами РНК являются конечным результатом экспрессии генов, и после ряда ферментативных модификаций они непосредственно используются в клеточных процессах. Это относится, прежде всего, к рибосомным и транспортным РНК (рРНК и тРНК), которые вместе составляют основную часть суммарной РНК клетки. К таким РНК принадлежат и малые ядерные РНК (мяРНК), участвующие в процессинге предшественников мРНК эукариот, РНК, входящие в состав ферментов, и природные антисмысловые РНК.

Синтез РНК происходит в результате сложной последовательности биохимических реакций, называемой *транскрипцией*. Появление русифицированного термина "мРНК" связано с тем, что на втором этапе реализации генетической информации, называемом *трансляцией*, последовательность нуклеотидов мРНК согласно генетическому коду однозначно определяет последовательность аминокислотных остатков синтезируемых белков, т.е. является матрицей, в соответствии с последовательностями нуклеотидов которой происходит соединение

аминокислотных остатков друг с другом в полипептидных цепях белков во время их биосинтеза. Таким образом, экспрессию генов определяют два глобальных молекулярно-генетических механизма: транскрипция генов и трансляция синтезированных мРНК рибосомами, которая завершается образованием полипептидных цепей, кодируемых генами. Однако процесс экспрессии генов не ограничивается их транскрипцией и трансляцией. Существенными моментами экспрессии генов являются *посттранскрипционные* и *посттрансляционные модификации* мРНК и белков, которые включают *процессинг* их предшественников (удаление избыточных последовательностей и другие ковалентные модификации последовательностей РНК и белков). Посттранскрипционные модификации предшественников мРНК обеспечивают подготовку мРНК к эффективной трансляции рибосомами и определяют продолжительность ее существования в цитоплазме. Посттрансляционные модификации белков также необходимы для их полноценного функционирования.

2.1. Транскрипция

В процессе транскрипции генов происходит биосинтез молекул РНК, комплементарных одной из цепей матричной ДНК, сопровождаемый полимеризацией четырех рибонуклеозидтрифосфатов (АТФ, ГТФ, СТР и УТР) с образованием 3'–5'-фосфодиэфирных связей и освобождением неорганического пирофосфата. Основными ферментами, осуществляющими транскрипцию, являются *ДНК-зависимые РНК-полимеразы*, которые функционируют с участием многочисленных *факторов транскрипции* – регуляторных белков, осуществляющих высокоспецифические белок–белковые и белково–нуклеиновые взаимодействия. Взаимодействия факторов транскрипции с регуляторными нуклеотидными последовательностями генов, друг с другом и с молекулами РНК-полимеразы необходимы для правильного узнавания транскрипционным комплексом регуляторных последовательностей в составе генов и приводят к повышению или понижению уровня транскрипции соответствующих последовательностей как ответ клеток на внешние или внутренние регуляторные сигналы. Благодаря факторам транскрипции и регуляторным последовательностям генов становится возможным

специфический синтез РНК и осуществляется регуляция экспрессии генов на уровне транскрипции.

2.1.1. ДНК-зависимые РНК-полимеразы

В соответствии с субъединичным составом РНК-полимеразы подразделяются на две группы. К первой группе относятся ферменты, состоящие только из одной субъединицы, среди них – РНК-полимеразы митохондрий и небольших бактериофагов, например SP6 и T7. Эти РНК-полимеразы транскрибируют небольшое число генов простых геномов, и для их функционирования не требуется сложных регуляторных воздействий. Вторую группу составляют сложно устроенные РНК-полимеразы бактерий и эукариот, которые представляют собой многосубъединичные белковые комплексы, транскрибирующие сотни и тысячи различных генов. Такие ферменты во время своего функционирования реагируют на многочисленные регуляторные сигналы, поступающие от регуляторных последовательностей нуклеотидов и белковых факторов. Не исключено, что общепринятое разделение РНК-полимераз по структурно-функциональному признаку является упрощением. Имеются данные, указывающие на то, что и просто устроенные фаговые РНК-полимеразы функционируют *in vivo* в комплексе с другими белками бактериальных клеток, которые могут существенно изменять их ферментативные свойства.

РНК-полимераза *E. coli*. Наиболее изученной из бактериальных ферментов является РНК-полимераза *E. coli*. Она осуществляет транскрипцию всех бактериальных генов. Фермент состоит из пяти субъединиц: β' - (молекулярная масса 165 кДа), β - (155 кДа), двух α - (35 кДа каждая) и σ - (чаще всего 70 кДа (σ^{70})). Комплекс из четырех субъединиц $\beta\beta'\alpha\alpha$, часто обозначаемый буквой E (enzyme), образует так называемый *минимальный (кор-) фермент* *E. coli*, который способен осуществлять все основные этапы транскрипции, за исключением правильной *инициации* (см. ниже). Для инициации транскрипции требуется присутствие определенной регуляторной σ -субъединицы, необходимой для распознавания РНК-полимеразой промоторов бактериальных генов, определяющей специфичность взаимодействия РНК-полимеразы с промоторами и, возможно, последующую изомеризацию

комплекса РНК-полимераза–промотор, необходимую для начала синтеза РНК. Полный фермент, включающий σ^{70} -субъединицу, часто называют *холоферментом* и обозначают $E\sigma^{70}$. РНК-полимераза $E\sigma^{70}$ способна транскрибировать большинство (но не все) генов *E. coli*. В частности, для транскрипции генов теплового шока, оперонов *gln* или *nif* требуется включение в состав полного фермента другой регуляторной субъединицы – σ^{54} (молекулярная масса 54 кДа) вместо σ^{70} с образованием фермента $E\sigma^{54}$. В настоящее время описано до десяти различных σ -факторов, объединение которых с минимальным ферментом дает возможность образующимся холоферментам узнавать разные промоторы. Все четыре субъединицы кор-фермента обеспечивают контакт РНК-полимеразы с промоторами. При этом β' -субъединица участвует в связывании фермента с ДНК, β -субъединица образует каталитический активный центр, а α -субъединицы обеспечивают правильное взаимодействие фермента с промоторами. Утверждения, заключенные в двух последних предложениях, нужно воспринимать с известной долей скепсиса. Данные такого рода обычно получают с использованием ферментов, у которых под действием мутаций изменены конкретные субъединицы, и если, например, мутация в гене α -субъединицы нарушает связывание РНК-полимеразы с ДНК, делаются соответствующие выводы. Такая методология (впрочем, одна из самых плодотворных среди существующих), к сожалению, напоминает известный способ локализации органа слуха у тараканов путем обрывания ног – поскольку тараканы без ног не реагируют на звуки убегающим, делается вывод, что они воспринимают звуковые сигналы ногами. Любая мутантная субъединица в составе олигомерного фермента может изменять его общую конформацию и придавать ферменту самые неожиданные свойства. Более прямым методом определения мест контакта макромолекул при белок–белковых и белково–нуклеиновых взаимодействиях является метод *поперечных сшивок* с использованием бифункциональных химических агентов. Такие химические соединения образуют ковалентные связи (поперечные сшивки) между близкорасположенными реакционноспособными группами. Однако сам факт наличия контакта между макромолекулами еще нельзя однозначно интерпретировать в пользу его функциональной значимости.

В отличие от эубактерий, которые, как уже упоминалось выше, при

транскрипции различных наборов генов используют разные σ -факторы, эукариоты для достижения тех же целей прибегают к другой стратегии – специализации молекул РНК-полимераз. В ядрах эукариот обнаружены по меньшей мере три специализированные формы РНК-полимераз. РНК-полимераза I осуществляет транскрипцию генов рибосомных РНК (рРНК), синтезируя в ядрышках предшественники 18S и 28S рРНК; РНК-полимераза II участвует в образовании мРНК, а РНК-полимераза III транскрибирует гены транспортных (тРНК), 5S и других низкомолекулярных РНК. Каждый из этих ферментов представляет собой многосубъединичный белковый комплекс, состоящий из двух больших (120–220 кДа) и 5–13 малых (10–100 кДа) субъединиц. Несколько малых субъединиц являются общими для разных форм РНК-полимераз. Большие же субъединицы гомологичны своими аминокислотными последовательностями участкам β - и β' -субъединиц эубактерий, что, возможно, отражает фундаментальное сходство в структуре и функционировании активных центров этих ферментов. Более того, аминокислотные последовательности α -субъединиц бактериальных РНК-полимераз, необходимые для их взаимодействия с большими субъединицами минимального фермента, имеют гомологи в третьей по размеру большой субъединице РНК-полимеразы II, а также в субъединице, общей у РНК-полимераз I и III. Несколько небольших субъединиц эукариотических РНК-полимераз, не имеющих аналогов у бактериальных ферментов, являются общими для всех РНК-полимераз, что может указывать на их одинаковые функции в транскрипции, осуществляемой соответствующими ферментами, и на их возможное участие в координации функционирования разных РНК-полимераз.

РНК-полимераза I эукариот (Pol I). Как и большинство других высокомолекулярных полипептидов, большие субъединицы РНК-полимераз содержат хорошо различимые структурные и функциональные домены: дискретные участки полипептидных цепей, несущие конкретную функциональную нагрузку. Клонирование генов соответствующих субъединиц и определение их первичной структуры позволили выявить эволюционно консервативные участки полипептидных цепей и провести мутационный анализ функциональной значимости их отдельных доменов. Для этой цели в

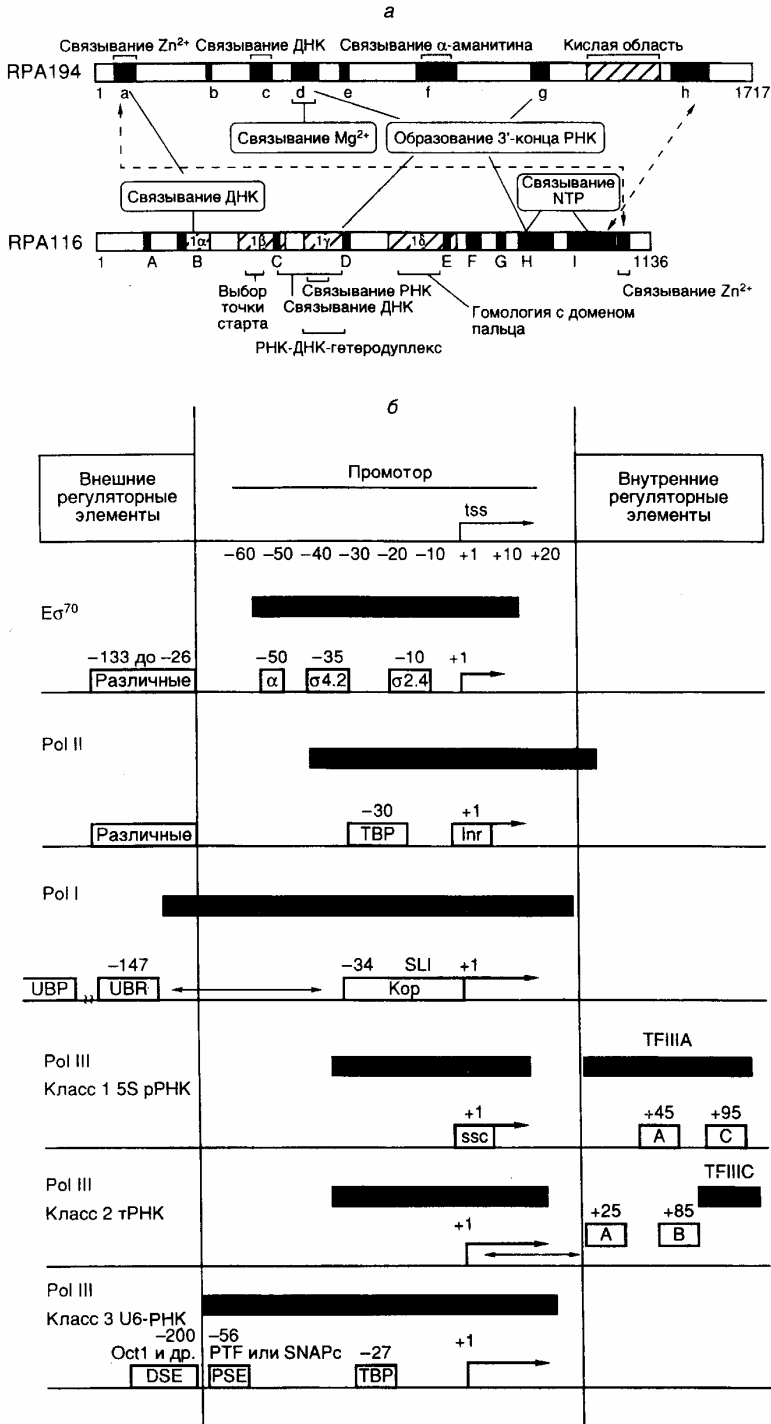


Рис. 1.4. Структурные и функциональные домены больших субъединиц эукариотической РНК-полимеразы I (а) и особенности структуры промоторов эубактерий и эукариот (б)

а – Полипептидные цепи двух больших субъединиц изображены в виде горизонтальных прямоугольников, в которых черным цветом и латинскими буквами отмечены участки, консервативные у большинства известных РНК-полимераз. Кислая область и участки $I\alpha$ - $I\delta$ характерны для РНК-полимераз I. Обозначены зоны полипептидных цепей, формирующие активный центр фермента и необходимые для выполнения соответствующих функций (например связывания Mg^{2+}). Пунктирные стрелки указывают на участки субъединиц, контактирующие друг с другом. б – Незаштрихованными прямоугольниками обозначены известные структурные элементы промоторов, необходимые для инициации или активации транскрипции. Внутри прямоугольников приведены названия факторов транскрипции, взаимодействующих с соответствующими элементами промоторов, а также названия сайтов или взаимодействующих с ними белков, находящихся над сайтами. Стрелки $\leftrightarrow$ обозначают фиксированные расстояния между элементами промоторов, а $\rightarrow$ – 5'-концевые части элонгируемых транскриптов. Черными прямоугольниками обозначены участки промоторов, защищаемые от действия ДНКазы I или других агентов $E\sigma^{70}$, а также эукариотическими транскрипционными комплексами, обеспечивающими базальный уровень транскрипции. tss – точка инициации транскрипции. Другие обозначения см. в тексте

полипептидных цепях с помощью направленного мутагенеза заменяли соответствующие аминокислоты и мутантные субъединицы использовали в сборке ферментов из отдельных субъединиц *in vitro* с последующим анализом свойств таких реконструированных ферментов. На рис. 1.4,а суммированы данные, полученные для двух самых больших субъединиц (RPA194 и RPA116) Pol I мышей, которые являются функциональными аналогами β' - и β -субъединиц РНК-полимеразы *E. coli*.

РНК-полимераза I эукариот является большим ферментом, построенным по меньшей мере из 11 субъединиц. Минимальный фермент Pol I содержит два обсуждавшихся выше больших полипептида с молекулярной массой 194 и 116 кДа, которые ассоциированы с несколькими малыми субъединицами (от трех до 14 в зависимости от метода очистки), молекулярные массы которых лежат в пределах 15–60 кДа. Третья по величине субъединица Pol I мышей с молекулярной массой 53 кДа, названная PAF53 (polymerase associated factor 53), играет важную роль в узнавании Pol I своих промоторов и, по-видимому, является структурным и функциональным аналогом белка RPA49 дрожжей. Pol I дрожжей в отсутствие субъединиц RPA49 и RPA35.5 (так называемая Pol I*) эффективно транскрибирует при низких концентрациях солей

искусственную матрицу poly[d(A-T)], но не нативную двухцепочечную ДНК. Полагают, что эти субъединицы необходимы для эффективного образования инициационных комплексов (см. ниже).

Используя антитела к отдельным субъединицам Pol I и последующую иммунопреципитацию, установили, что в клетке, по крайней мере, часть Pol I находится в составе больших комплексов, с которыми ассоциированы факторы транскрипции. Пять компонентов такого холофермента Pol I изучены в настоящее время наиболее детально.

Мышиный фактор TIF-IB (Pol I-specific transcription initiation factor B), известный также, как фактор D, обеспечивает Pol I селективность в отношении промоторов генов рПНК (рДНК). Аналогичный белок у человека назван hSL1, у крыс – rSL1 и у *X. laevis* – Rib 1. Взаимодействие фактора TIF-IB/SL1 с промотором рДНК обеспечивает связь холофермента Pol I с промотором и сборку прединициационного комплекса. Фактор TIF-IB/SL1 состоит из четырех субъединиц, одна из которых является основным фактором транскрипции TBP, необходимым для функционирования РНК-полимераз всех трех классов. (Подробнее об основных факторах транскрипции см. следующий раздел 2.1.3.) Три других субъединицы с молекулярными массами 110, 63 и 48 кДа представляют собой разные TBP-ассоциированные факторы TAF_I, индивидуально и специфически взаимодействующие с TBP, а также друг с другом, образуя прочный комплекс. В составе комплекса TAF_I48 обеспечивает контакт TIF-IB/SL1 с фактором UBF (см. ниже), а TAF_I63 и TAF_I110 участвуют в распознавании промотора. Факторы TAF_I не обнаруживают гомологии с соответствующими факторами TAF_{II}, специфичными в отношении Pol II. Более того, первый из связавшихся с TBP факторов TAF_I предотвращает взаимодействие с TBP факторов TAF_{II} (и наоборот), что делает невозможным образование непродуктивных химерных комплексов. Одновременно взаимодействие TAF_I48 с TBP изменяет ДНК-связывающие свойства последнего, после чего тот перестает узнавать TATA-бокс – характерный структурный элемент Pol II-промоторов, и, следовательно, теряет способность обеспечивать инициацию транскрипции Pol II.

Другой белок, входящий в состав холофермента Pol I, UBF (upstream binding factor) высоко консервативен у разных видов животных. UBF является членом семейства факторов транскрипции, содержащих ДНК-связывающий

HMG-домен (high mobility group domain) – основную последовательность из 80 аминокислот. С помощью ЯМР-спектроскопии установлено, что полипептидная цепь HMG-домена организована в три α -спирали, расположенные в виде буквы L, которые формируют три ДНК-связывающих поверхности с внешней стороны L. В клетке UBF присутствуют в двух формах – UBF1 и UBF2 с молекулярными массами 97 и 95 кДа, которые образуются в результате альтернативного сплайсинга. UBF1 содержит пять HMG-доменов, фланкированных N-концевым димеризующим мотивом и короткой кислой C-концевой последовательностью. Интересно, что соседние HMG-домены одного и того же UBF обладают гораздо меньшей гомологией, чем соответствующие домены UBF разных видов (например шпорцевой лягушки и человека). Полагают, что каждый HMG-домен обеспечивает особую, эволюционно консервативную функцию молекулы UBF. Такими функциями могут быть распознавание специфических последовательностей ДНК, создание молекулярных интерфейсов для белок–белковых взаимодействий между Pol I и TIF-IB/SL1, а также различными репрессорами и активаторами транскрипции рДНК. C-Концевая последовательность UBF содержит несколько фосфорилируемых остатков Ser и необходима для активации транскрипции рДНК. Одной из основных характеристик белков, содержащих HMG-бокс, является их способность изгибать молекулу ДНК и прочно связываться с ее крестообразными структурами. Всеми этими свойствами обладает UBF, и они детально исследованы.

Белок CPBF (core promoter binding factor), выделенный из асцитных клеток аденокарциномы молочных желез крыс, специфически взаимодействует с коровым участком промотора рДНК (о структуре промотора см. в разделе 2.1.2.). CPBF, прочно взаимодействующий с Pol I, состоит из двух субъединиц USF1 и USF2 с молекулярными массами 44 и 39 кДа соответственно. Гомодимеры USF1 и USF2 являются сильными ингибиторами транскрипции Pol I, тогда как гетеродимеры USF1/USF2 стимулируют транскрипцию *in vitro*. Полагают, что CPBF участвует в регуляции транскрипции Pol I *in vivo*.

TIF-IA – другой компонент холофермента Pol I, также участвует в регуляции синтеза рРНК этим ферментом. В его отсутствие инициационный комплекс не может образовывать первой фосфодиэфирной связи, а следовательно, и инициировать синтез РНК. TIF-IA освобождается после

инициации транскрипции и может вновь входить в состав собирающихся прединициационных комплексов. По этим и ряду других критериев TIF-IA рассматривают в качестве функционального аналога бактериального фактора σ^{70} . TIF-IA является мономерным глобулярным белком с молекулярной массой 70–80 кДа. Активность этого фактора или его внутриклеточное содержание уменьшается при подавлении синтеза белка, истощении сыворотки или дифференцировке клеток и возрастает в ответ на митогенные стимулы, что коррелирует с подавлением или стимуляцией синтеза рРНК.

Хроматографически и с помощью иммунопреципитации было установлено, что жизненно важный фактор TIF-IC в растворе ассоциирован с Pol I. Этот фактор необходим как для сборки инициационных комплексов, так и образования первой фосфодиэфирной связи. Его присутствие предотвращает неспецифическую инициацию транскрипции и ее преждевременную терминацию, что проявляется в образовании гомогенных транскриптов правильной длины. По этим критериям фактор TIF-IC рассматривают в качестве функционального аналога TFIIIF (RAP30/74) Pol II (см. ниже).

РНК-полимераза II (Pol II). Pol II человека содержит более 10 субъединиц, которые трудно назвать субъединицами в обычном смысле из-за слабой ассоциации друг с другом. Некоторые из них принадлежат к *основным факторам транскрипции* (GTFs – general transcription factors). Вообще же понятие холофермента Pol II эукариот не является устоявшимся. Лишь недавно в лабораториях Р.Янга и Р.Корнберга было установлено, что некоторые основные факторы транскрипции уже находятся в комплексе с РНК-полимеразой до ее включения в прединициационный комплекс. По мнению Янга, которое становится все более обоснованным, в состав холофермента Pol II дрожжей входят по меньшей мере 14 белков и белковых комплексов, перечисленных в табл. I.3.

Таблица I.3

**Характеристики белковых компонентов холофермента
РНК-полимеразы II дрожжей**

Компонент	Характеристика
Pol II	РНК-Полимеразная активность, взаимодействует с множеством общих и тканеспецифических факторов транскрипции, участвует в выборе точки инициации транскрипции
TFIIB	Связывает Pol II и TBP на промоторе, участвует в выборе точки инициации транскрипции
TFIIF	Взаимодействует с Pol II, стимулирует элонгацию транскрипции Pol II, компонент субкомплекса SRB/медиатор
TFIIN	Активность ДНК-зависимой АТРаза, ДНК-хеликазная активность, обладает активностью CTD-киназы
SRB2, SRB5	Участвуют в образовании инициационного комплекса, стимулируют базальный и индуцированный синтез РНК, взаимодействуют с TBP, компоненты субкомплекса SRB/медиатор
GAL11/SPT13	Участвуют в образовании инициационного комплекса, стимулируют базальный и индуцированный синтез РНК, компоненты субкомплекса SRB/медиатор, предположительно взаимодействуют с активаторами транскрипции
SUG1	Компонент субкомплекса SRB/медиатор, предположительно взаимодействует с активаторами транскрипции
SRB4, SRB6, SRB7, SRB8, SRB9, SRB10, SRB11	Компоненты субкомплекса SRB/медиатор, предположительно взаимодействуют с CTD-доменом Pol II

Среди факторов, которые отличаются от основных факторов транскрипции, но играют особую роль в транскрипции у дрожжей, следует отметить так называемые SRB-белки (suppressor of RNA polymerase B). Последние являются неотъемлемой частью холофермента Pol II и относятся к коактиваторам транскрипции. SRB-Белки идентифицированы с помощью генетических методов при отборе мутаций у дрожжей, которые супрессировали условно-летальные мутации в CTD-домене Pol II (С-концевом домене большой субъединицы РНК-полимеразы II). Такой генетический скрининг привел к открытию девяти генов *SRB*. С помощью тех же мутантов установлено, что SRB-белки необходимы для осуществления инициации транскрипции РНК-полимеразой *in vivo*. Вскоре было подтверждено физическое и функциональное взаимодействие между SRB-белками и CTD-доменом Pol II, которые, как оказалось, образуют с РНК-полимеразой прочный комплекс.

Транскрипция с участием холофермента Pol II стимулируется активатором GAL4–VP16, что не происходит в присутствии только одних очищенных основных факторов транскрипции и Pol II. На этом основании был сделан вывод, что истинный холофермент Pol II содержит дополнительные компоненты, которые позволяют ему реагировать на действие белков-активаторов транскрипции. Более того, показано, что антитела к CTD-домену вызывают диссоциацию многосубъединичного комплекса, содержащего Pol II, SRB-белки, TFIIF, GAL11/SPT13, SUG1 и еще ~10 белков, часть из которых относится к классу белков-коактиваторов транскрипции. Этот комплекс белков получил название *медиаторного комплекса*. Он оказался необходимым для осуществления функциональной связи между Pol II и белками-активаторами транскрипции (подробнее о коактиваторах транскрипции см. раздел 3.2.2).

Субъединичное строение РНК-полимераз разного происхождения, вероятно, отражает их функциональную роль в акте транскрипции. Действительно, все РНК-полимеразы простого строения транскрибируют узкоограниченный круг генов или небольшие части генома, что имеет место, например при синтезе РНК-затравок для фрагментов Оказаки в процессе репликации ДНК у бактерий. Промоторы, узнаваемые РНК-полимеразами простого строения, не отличаются разнообразием и обладают простой структурой. Показательно, что при сложном строении генома четных Т-фагов, в процессе развития которых происходит многократное переключение

транскрипции с одних групп генов на другие, используется сложная РНК-полимераза бактерии-хозяина, а не индуцируется простой фермент, как это имеет место у бактериофага Т7.

РНК-полимеразы бактерий и эукариот должны, во-первых, узнавать разные промоторы, во-вторых, реагировать на различные белки-регуляторы и, в-третьих, изменять специфичность узнавания последовательностей нуклеотидов матричных ДНК под действием разнообразных белковых эффекторов. Отсюда следует, что у живых организмов, начиная с бактерий, возникает потребность в РНК-полимеразах сложной структуры, способных осуществлять обширную программу реализации генетической информации. Вероятно, поэтому наблюдается иерархия в степени сложности строения указанных ферментов, которая достигает верхнего предела в случае РНК-полимераз эукариот.

Тем не менее, элементарные акты основных этапов транскрипции (см. ниже) обеспечиваются молекулами РНК-полимераз простого строения, такими как у фагов Т7, митохондрий и других объектов. Эти ферменты, по мнению Р.Б. Хесина, можно рассматривать в качестве эволюционных предшественников сложных олигомерных РНК-полимераз, способных самостоятельно осуществлять все основные функции в процессе транскрипции. Действительно, у олигомерных РНК-полимераз, как и у большинства сложноустроенных ферментов, по-видимому, только одна субъединица (β – у РНК полимераз зубактерий) является собственно каталитической, а остальные, возможно, выполняют функции регуляторных.

2.1.2. Единицы транскрипции (транскриптоны)

Синтез РНК молекулами РНК-полимераз *in vivo* начинается в определенных местах ДНК, называемых *промоторами*, и завершается на особых регуляторных последовательностях – *терминаторах*. Последовательности нуклеотидов ДНК, заключенные между промоторами и терминаторами, называют *транскрипционными единицами*, или *транскриптонами*. В пределах каждого транскриптона транскрибируется только одна цепь ДНК, которая получила название *значащей* или *матричной*. Термины "транскрипционная единица" или "транскриптон" по смыслу близки

термину "ген", но они не всегда совпадают. Так, транскрипционные единицы прокариот, как правило, заключают в себе генетическую информацию нескольких генов и называются *оперонами*. Продуктами транскрипции оперонов являются *полицистронные* мРНК, в результате трансляции которых рибосомами образуется несколько белков. Белки, кодируемые полицистронными мРНК, обычно функционально связаны друг с другом и обеспечивают протекание какого-либо метаболического процесса, например биосинтеза определенной аминокислоты или утилизацию углеводов в качестве источника углерода. Организация генов в виде оперонов облегчает координированную регуляцию их экспрессии на уровне транскрипции. Согласованная регуляция транскрипции (и других этапов экспрессии) многих генов, не образующих одного оперона, чаще всего осуществляется специфическими белками-регуляторами, которые взаимодействуют с гомологичными регуляторными нуклеотидными последовательностями, маркирующими гены данной группы.

Промоторы зубактерий. В 1964 г. Ф. Жакобом, А. Уллманом и Ж. Моно была впервые сформулирована концепция промотора как особого регуляторного элемента в составе *lac*-оперона *E. coli*. Вскоре после этого были описаны промоторные мутации и высказано предположение, что промоторные участки оперона должны распознаваться молекулами РНК-полимеразы, а также заключать в себе сайт инициации транскрипции. В соответствии с генетическими критериями промотором может называться последовательность нуклеотидов, мутации в которой характеризуются следующими свойствами: 1) оказывают влияние на транскрипцию всех генов оперона; 2) действуют в *цис*-, но не *транс*-положении по отношению к генам, контролируемым мутантными промоторами; 3) локализуются вблизи одного из концов оперона, однако отличаются от мутаций в структурных частях генов тем, что не оказывают влияния на активность кодируемых этими генами ферментов. В соответствии с биохимическими критериями промотор представляет собой последовательность нуклеотидов, обеспечивающую базальный (но не максимальный) уровень транскрипции соответствующего транскрипта. Он является той минимальной последовательностью, которая специфически распознается холоферментом РНК-полимеразы среди случайных последовательностей нуклеотидов, что обнаруживается по образованию в

определенных условиях стабильных комплексов фермент–ДНК.

Структура обобщенного минимального промотора для холофермента РНК-полимеразы ($E\sigma^{70}$) эубактерий схематически представлена на рис. 1.4, б. Такие промоторы содержат две канонические последовательности: область -35, обычно расположенную между нуклеотидами в положениях -30 и -35, и область -10, локализованную между нуклеотидами -7 и -10 (знак "минус" указывает на то, что нуклеотиды находятся перед первым транскрибируемым нуклеотидом, которому соответствует положение +1). Обе эти последовательности специфически взаимодействуют с двумя участками полипептидной цепи σ^{70} , называемыми $\sigma 4.2$ и $\sigma 2.4$. Расстояние между обсуждаемыми последовательностями оказывает влияние на активность промотора. Длина отрезка в 17 п.о. является оптимальной. Среди других последовательностей, оказывающих влияние на активность промоторов, но не являющихся универсальными, следует упомянуть так называемый UP-элемент – АТ-богатую последовательность, центр которой находится в положении -52 у промотора P1 гена *rrnB*, кодирующего рРНК. С этим участком взаимодействует α -субъединица РНК-полимеразы. У нескольких других промоторов был обнаружен TG-элемент (центр – в положении -15 или -14), мутации в котором снижают активность этих промоторов.

Влияние мутаций в областях -10 и -35 на активность промотора, как правило, хорошо коррелирует с тем, насколько новая мутантная последовательность соответствует канонической. Однако в ряде случаев наблюдаются отклонения от этого правила, что позволило идентифицировать последовательности, важные для функционирования этих промоторов *in vivo*, с центрами в положениях -43 (область USR – upstream region), а также между нуклеотидами в положениях -1 и +20 (область DSR – downstream region). Эти последовательности не требуются для узнавания промотора РНК-полимеразой, однако облегчают процесс освобождения промотора после инициации транскрипции (см. ниже). Все описанные последовательности обеспечивают базальный уровень функционирования бактериальных промоторов. Однако активность промоторов может изменяться под действием многочисленных регуляторных факторов, взаимодействующих с другими регуляторными последовательностями, расположенными перед такими промоторами. Эти

аспекты регуляции активности промоторов будут подробно рассмотрены в разделе 3.1.

Промоторы эукариот. Промоторы эубактерий были определены выше как минимальный набор последовательностей нуклеотидов, необходимых для их специфического распознавания холоферментом РНК-полимеразы и начала (инициации) синтеза РНК. Распространение этого определения на промоторы эукариот встречает трудности из-за невозможности в настоящее время исчерпывающе описать эукариотический функциональный аналог бактериального холофермента $E\sigma$.

Промоторы РНК-полимеразы II. Промоторы, узнаваемые РНК-полимеразой II, содержат три различных семейства регуляторных последовательностей ДНК. Последовательности первого семейства включают так называемые *кóровые*, или базальные элементы промотора, расположенные вблизи точки инициации транскрипции (см. рис. 1.4,б). В настоящее время известны два класса базальных элементов: ТАТА-последовательность, расположенная за 25–30 нуклеотидов до точки инициации (каноническая последовательность – $TATAa/tAa/t$), и так называемый *инициатор* (Inr), последовательность которого обогащена пиримидинами (каноническая последовательность – $PuPuA_{+1}NT/APuPu$). В этих обозначениях строчные буквы соответствуют нуклеотидам, которые присутствуют не во всех последовательностях, Pu – пиримидиновые нуклеотиды инициатора, N – любой нуклеотид. Подстрочная цифра указывает на то, что с данного нуклеотида начинается транскрипция (*точка инициации транскрипции*). Элементы ТАТА-последовательности и инициатор необходимы для сборки ДНК-белкового инициационного комплекса и распознаются основными факторами транскрипции. Промоторы РНК-полимеразы II содержат один или оба регуляторных элемента или же не имеют их вообще. При этом оба элемента могут функционировать независимо друг от друга или же в их действии наблюдается синергизм.

К двум другим классам *цис*-регуляторных промоторных элементов у эукариот относятся последовательности, расположенные вблизи промотора (от 50 до нескольких сотен пар оснований перед точкой инициации), а также дистальные элементы (энхансеры и сайленсеры), расстояние которых от промотора может превышать 60 т.п.о. Оба класса таких последовательностей

содержат сайты связывания регуляторных белков, модулирующих транскрипцию.

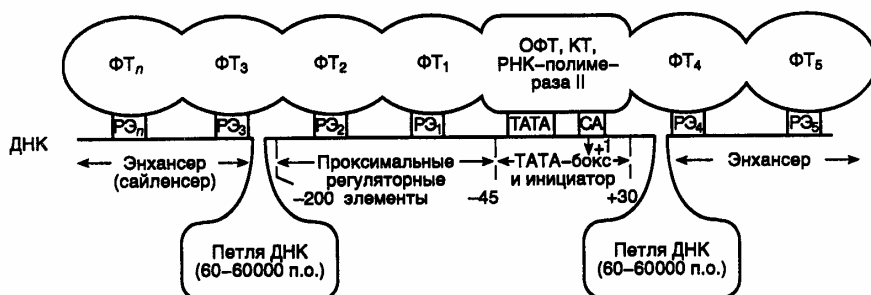


Рис. 1.5. Межмолекулярные взаимодействия на промоторе РНК-полимеразы II

РЭ₁–РЭ_n – последовательности нуклеотидов регуляторных элементов промотора, ФТ₁–ФТ_n – взаимодействующие с ними регуляторные факторы транскрипции, ОФТ – основные факторы транскрипции, КТ – коактиваторы транскрипции, +1 – точка инициации транскрипции

Обобщенная схема промотора РНК-полимеразы II в составе инициационного комплекса представлена на рис. 1.5. На ней отображены многочисленные белок–белковые и белково–нуклеиновые взаимодействия при инициации транскрипции. На этой схеме также указано наличие во многих промоторах консервативной последовательности СА, которая находится непосредственно перед нуклеотидом А₊₁. У некоторых промоторов, в частности ассоциированных с генами домашнего хозяйства, может отсутствовать явно выраженная ТАТА-последовательность. Проксимальные и дистальные регуляторные элементы промоторов построены из коротких транскрипционных элементов длиной в 10–15 п.о., с которыми непосредственно взаимодействуют факторы транскрипции. Проксимальные регуляторные элементы, как правило, имеют простую структуру, включающую один или несколько транскрипционных элементов (ТЭ). В то же время энхансеры устроены более сложно. Следует заметить, что в активном промоторе дистальные и проксимальные регуляторные элементы сближены друг с другом, в результате чего происходит "выпетливание" разделяющих их протяженных последовательностей ДНК.

Промоторы РНК-полимеразы III. Как уже упоминалось выше, гены, транскрибируемые РНК-полимеразой III, можно разделить на три основных класса (см. рис. 1.4,б). К 1-му классу относят гены 5S рРНК, которые содержат

внутреннюю регуляторную область (internal control region – ICR). Три элемента ICR, с которыми взаимодействует основной фактор транскрипции TFIIIA, локализованы между нуклеотидами +50 и +64. Комплекс TFIIIA–ICR стабилизируется после взаимодействия с TFIIIA фактора TFIIIC, хотя для последнего в промоторе отсутствует специфический участок связывания (подробнее о факторах транскрипции см. в разделах 2.1.3 и 3.2). Мутации, локализованные в окрестностях точки инициации транскрипции промоторов классов 1 и 2, оказывают сильное влияние на транскрипцию этих генов, что указывает на возможное присутствие в этих промоторах дополнительных регуляторных элементов.

Гены тРНК относят ко второму классу. Для промоторов этих генов характерно наличие внутренних блоков регуляторных последовательностей А и В. Блок А локализован между нуклеотидами +10 и +20, тогда как блок В удален от него на расстояние 30–60 п.о. по направлению к 3'-концу гена. Эти регуляторные элементы узнаются непосредственно фактором TFIIIC, который образует с ними очень прочный комплекс. По крайней мере, четыре полипептидные цепи, входящие в состав комплекса TFIIIC–промотор, контактируют друг с другом и покрывают участок ДНК между положениями –25 и +75. Канонический регуляторный элемент в окрестностях точки инициации транскрипции был локализован путем анализа последовательностей нуклеотидов в соответствующих генах тРНК.

Гены третьего класса кодируют малые ядерные РНК, например U6-РНК у высших эукариот, и содержат только регуляторные элементы, расположенные перед точкой инициации транскрипции. У большинства генов этого класса регуляторные элементы состоят из TATA-последовательности с центром в положении –30 и из проксимального элемента PSE (proximal sequence element) с центром в положении –60. Несмотря на наличие TATA-элемента, обычно определяющего положение точки инициации транскрипции, в данном случае именно PSE фиксирует положение этой точки. Аналогичную функцию у промоторов РНК-полимеразы II выполняет TATA-последовательность. Состав белковых компонентов, образующих комплекс с PSE, в настоящее время до конца не выяснен. Обнаружены два различных PSE-связывающих белковых комплекса, названные SNAP и PTF. Активность промоторов третьего класса зависит от наличия TATA-связывающего белка (TBP). Белок TBP, входящий в

состав фактора транскрипции TFIIIB, взаимодействует непосредственно с TATA-элементом промотора. Поскольку для инициации транскрипции на этих промоторах требуется наличие PSE-элемента, предполагают, что PSE-связывающий фактор взаимодействует с TBP-компонентом фактора TFIIIB. Интересно, что промотор гена U6-РНК животных, узнаваемый РНК-полимеразой III, отличается от промотора гена U2-РНК, с которым взаимодействует РНК-полимераза II, только отсутствием TATA-элемента U2-промотора. У растений оба этих промотора содержат такие регуляторные элементы, а изменение расстояния между TATA-элементом и расположенным перед ним регуляторным элементом превращает промотор U6 из Pol III-зависимого в Pol II-зависимый. Несмотря на наличие у промоторов этих РНК-полимераз характерных черт, описанных выше, механизмы, по которым РНК-полимеразы II и III различают свои промоторы, остаются невыясненными.

Промоторы РНК-полимеразы I. Несмотря на то, что участки генома, кодирующие рРНК, относятся к одним из наиболее консервативных, регуляторные последовательности, обеспечивающие инициацию транскрипции у разных видов, сильно дивергированы. Одно из объяснений этого явления основывается на факте вхождения рДНК в мультигенное семейство (гены рРНК многократно повторены в геноме). В этой связи мутации в регуляторной области одного гена, которые благоприятно сказываются на его экспрессии, могут быстро распространиться в семействе последовательностей благодаря генной конверсии, неравному кроссинговеру и/или последовательным актам их эксцизии и реинтеграции в геном. В соответствии с этой интерпретацией, промоторные последовательности рДНК обнаруживают мало гомологии даже у таких близких видов, как человек и мышь. Тем не менее, общий план строения промоторов рДНК практически не различается у всех исследованных эукариот. Промоторы РНК-полимеразы I содержат два коровых регуляторных элемента (CPE – core promoter element), локализованных между положениями –75 и –50, а также –30 и +1 (см. рис. 1.4, б). Основным промотор лишь слабо связывает TBP-содержащий белковый комплекс SL1 (у людей и крыс), в состав которого, кроме того, входят еще три видоспецифических фактора транскрипции (TBP associated factor – TAF), специфичных в отношении РНК-полимеразы I. Сборка прединициационного комплекса (см. раздел 2.1.3) в этом случае не зависит от специфического взаимодействия TBP и ДНК, так как промоторы РНК-

полимеразы I не содержат ТАТА-последовательности. Как уже упоминалось выше, TAF₆₃ и TAF₁₁₀, входящие в состав комплекса SL1, являются ДНК-распознающими компонентами комплекса. С помощью футпринтинга (см. раздел 7.6.4) было установлено, что комплекс SL1 остается связанным с промотором на протяжении нескольких раундов инициации синтеза РНК РНК-полимеразой I. Стабильное взаимодействие комплекса SL1 с промотором требует участия дополнительного ДНК-связывающего белка UBF, который сам по себе обладает способностью слабо связываться с удаленной регуляторной последовательностью UCE (upstream control element), расположенной между нуклеотидами в положениях –90 и –150. Синергизм в стимулирующем действии SL1 и UBF на транскрипцию РНК-полимеразой I может быть обусловлен наличием непосредственных контактов между этими двумя факторами. Другим регуляторным элементом, удаленным от точки инициации транскрипции промотора Pol I, является энхансер, построенный из коротких повторяющихся последовательностей.

Последовательность CPE обеспечивает специфичность транскрипции генов рРНК, и ее достаточно для инициации. В этой связи последовательность UCE и энхансер не рассматриваются в настоящее время как части промотора РНК-полимеразы I. Кроме того, соответствующие промоторы простейших, грибов и растений последовательность UCE не содержат, а белок UBF не является необходимым компонентом бесклеточных систем транскрипции. Как будет подробно рассмотрено ниже, образование прединициационного комплекса РНК-полимеразой II требует участия фактора TFIIB. Аналогичный в функциональном отношении белковый фактор BRF также необходим для инициации транскрипции РНК-полимеразой III. Соответствующий белковый компонент РНК-полимеразы I еще не идентифицирован. Помимо фактора, связывающего коровую последовательность – CPBF, в инициации синтеза рРНК принимает участие фактор, связывающий энхансер 1, – E1BF (enhancer 1-binding factor), который не входит в состав холофермента Pol I. Оба фактора заметно повышают базальный уровень синтеза рРНК в бесклеточных системах транскрипции.

2.1.3. Этапы транскрипции

Процесс транскрипции в настоящее время принято подразделять на 4 основные стадии: 1) связывание молекул РНК-полимеразы с ДНК и распознавание промотора; 2) *инициация*; 3) *элонгация*; 4) *терминация*. Как будет видно из дальнейшего изложения, три последних этапа характерны для биосинтеза большинства других макромолекул клетки, особенно для тех из них, синтез которых является матричным, в частности белков. После связывания с ДНК молекулы РНК-полимеразы осуществляют поиск промоторов, на которых происходит формирование *инициационных комплексов*. Начальная стадия инициации транскрипции завершается образованием нескольких первых фосфодиэфирных связей в молекуле вновь синтезируемой РНК, после чего транскрипция переходит в стадию элонгации – последовательного удлинения синтезируемых молекул РНК. Стадия элонгации заканчивается по достижении молекулами РНК-полимераз специальных регуляторных последовательностей ДНК, называемых терминаторами транскрипции, после чего происходит освобождение синтезированных молекул РНК и РНК-полимераз из транскрипционных комплексов. Освободившиеся молекулы РНК-полимераз приобретают способность вступать в новый цикл транскрипции. Следует помнить, что четкое разделение единого процесса транскрипции на отдельные стадии условно; оно используется главным образом для удобства описания механизмов биосинтеза РНК и является упрощенной моделью. Основные этапы транскрипции, а также дальнейшие пути реализации генетической информации схематически представлены на рис. 1.6.

В обычных условиях холофермент РНК-полимераз эубактерий для инициации транскрипции не требует дополнительных факторов. В отличие от этого для точной инициации транскрипции РНК-полимеразой II требуется наличие, кроме ее субъединиц, еще и основных факторов транскрипции. Синтез РНК, который не зависит от присутствия регуляторных молекул, получил название *базальной транскрипции*. Транскрипция в клетках является регулируемым процессом, который, как уже упоминалось выше, требует участия белков-активаторов или репрессоров. Белок-активатор (тканеспецифический фактор транскрипции) взаимодействует с регуляторными последовательностями ДНК и активирует синтез РНК. Такая транскрипция

получила название *индуцированной*, или *активированной*. Следовательно, базальная транскрипция не может происходить *in vivo*, и этот термин используется только при описании результатов исследований синтеза РНК *in vitro*, в бесклеточных системах транскрипции.

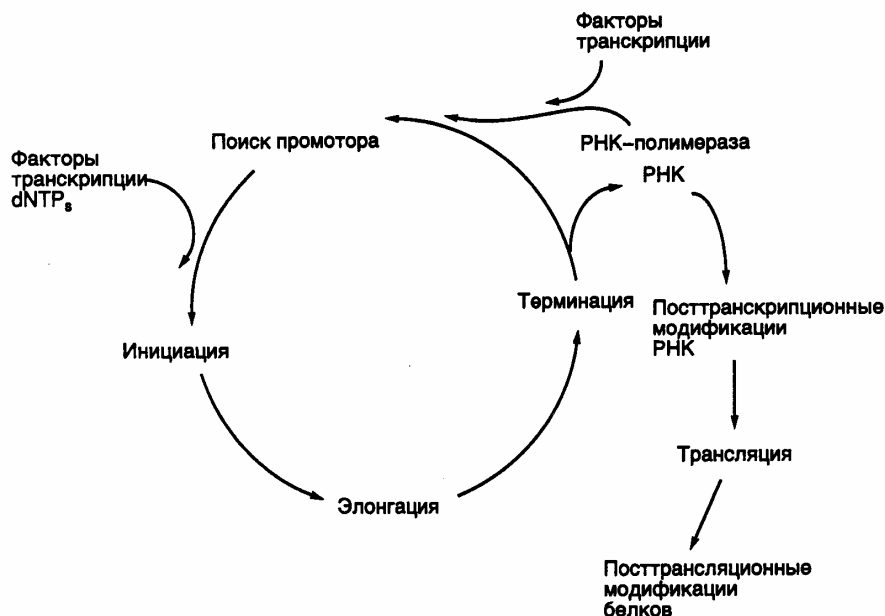


Рис. 1.6. Цикл транскрипции и последующие пути реализации генетической информации

Связывание молекул РНК-полимеразы с ДНК и поиск промоторов.

Механизм поиска промоторов на ДНК молекулами РНК-полимеразы в настоящее время до конца не выяснен. Принято считать, что после первоначального непрочного связывания с ДНК в случайных местах молекулы РНК-полимеразы перемещаются вдоль двойной спирали ДНК до тех пор, пока не обнаруживают последовательности нуклеотидов промоторов, на которых взаимодействие фермента с ДНК становится более прочным. Во время движения молекулы РНК-полимеразы могут периодически отделяться от ДНК и связываться с ней на новом месте, что ускоряет процесс поиска промоторов. Как уже упоминалось выше, в связывании с ДНК участвует β -субъединица РНК-полимеразы *E. coli*, а α - и особенно σ -субъединицы необходимы для специфического распознавания промоторов. Установлено, что холофермент РНК-полимеразы *E. coli* (минимальный фермент, содержащий σ -субъединицу)

закрывает в области промотора участок ДНК длиной ~50 п.о. При этом α -субъединицы контактируют с ДНК в области -35-го нуклеотида промотора.

Инициация транскрипции. Инициация транскрипции начинается со сборки на промоторе прединициационного комплекса, в состав которого входят молекулы РНК-полимеразы и матричной ДНК. Если в случае РНК-полимеразы *E. coli* и других прокариот для осуществления этого процесса нет необходимости в присутствии других белковых факторов, то механизм сборки инициационного комплекса с участием РНК-полимеразы II носит более сложный характер. В настоящее время существуют две модели инициации транскрипции РНК-полимеразой II. В соответствии с одной из них на промоторе происходит постепенная (ступенчатая) сборка инициационного комплекса из отдельных компонентов. Другая модель акцентирует внимание на то, что Pol II может входить в состав инициационного комплекса в виде холофермента, состоящего из многих субъединиц. Хотя вторая модель становится доминирующей, ниже будет подробнее рассмотрена первая модель, более наглядно описывающая процесс инициации транскрипции у эукариот.

Сборка такого комплекса начинается с последовательного связывания с промотором *основных факторов транскрипции* (табл. I.4). Обычно факторами транскрипции называют белки или белковые комплексы, непосредственно не участвующие в каталитическом акте образования РНК, но необходимые для прохождения основных этапов транскрипции и ее регуляции. По функциональному признаку принято различать три класса факторов транскрипции. К первому классу относятся основные факторы транскрипции, обеспечивающие нерегулируемый базальный уровень транскрипции и функционирующие в клетках всех типов.

Таблица I.4

**Субъединичный состав и характеристика основных факторов
транскрипции (GTF) РНК-полимеразы II человека**

Фактор (GTF)	Молекулярные массы субъединиц (кДа) и их обозначение	Характеристика
TFIIA	37 (α) 19 (β) 13 (γ)	Требуются для активации транскрипции, стабилизируют взаимодействие TBP с ТАТА- боксом, а также TAF с ДНК, необходимы для активации некоторых промоторов
TFIIB	35	Стабилизирует комплекс Pol II/TFIIF на промоторе, обеспечивает выбор точки инициации транскрипции, претерпевает конформационные изменения под действием белков-активаторов
TFIID	38 (TBP)	Узнает ТАТА-последовательность, является мишенью для белков-трансактиваторов транскрипции
	250 (TAFII250)	Сериновая протеинкиназа, узнает HMG-бокс
	150 (TAFII150)	Узнает 3'-концевые области промотора
	135 (TAFII135)	
	95 (TAFII95)	
	80 (TAFII80)	Гомолог гистона H4
	55 (TAFII55)	
	31 (TAFII31)	Гомолог гистона H3
	28 (TAFII28)	
	20 (TAFII20)	Гомолог гистона H2B

Таблица 1.4 (окончание)

Фактор (GTF)	Молекулярные массы субъединиц (кДа) и их обозначение	Характеристика
TFIIE	56	Ассоциированы с Pol II (другое название – RAP), обладают Zn-связывающими доменами, обе субъединицы участвуют в плавлении промотора, обеспечивают входение TFIIN в прединициационный комплекс
	34	
TFIIF	58 (RAP74)	Стимулирует элонгацию, фосфорилируется in vivo Гомолог σ -фактора РНК-полимеразы E. coli, предотвращает неспецифическую инициацию транскрипции
	26 (RAP30)	
TFIIN	89 (ERCC3)	3'→5'-Хеликаза, необходимая для транскрипции; участвует в эксцизионной репарации ДНК
	80 (ERCC2)	
	62 (p62)	Участвует в эксцизионной репарации ДНК
	44 (hSSL1)	
		Содержит Zn-связывающий домен, участвует в эксцизионной репарации ДНК
	40 (cdk7)	
	37 (циклин H)	Содержит Zn-связывающий домен, гомологичный hSSL1
	34 (p34)	
TFII I	32 (MAT-1)	Содержит Zn-связывающий домен; фактор сборки CAK
	120	
		Связывает INR-элемент (инициатор), облегчает связывание TBP

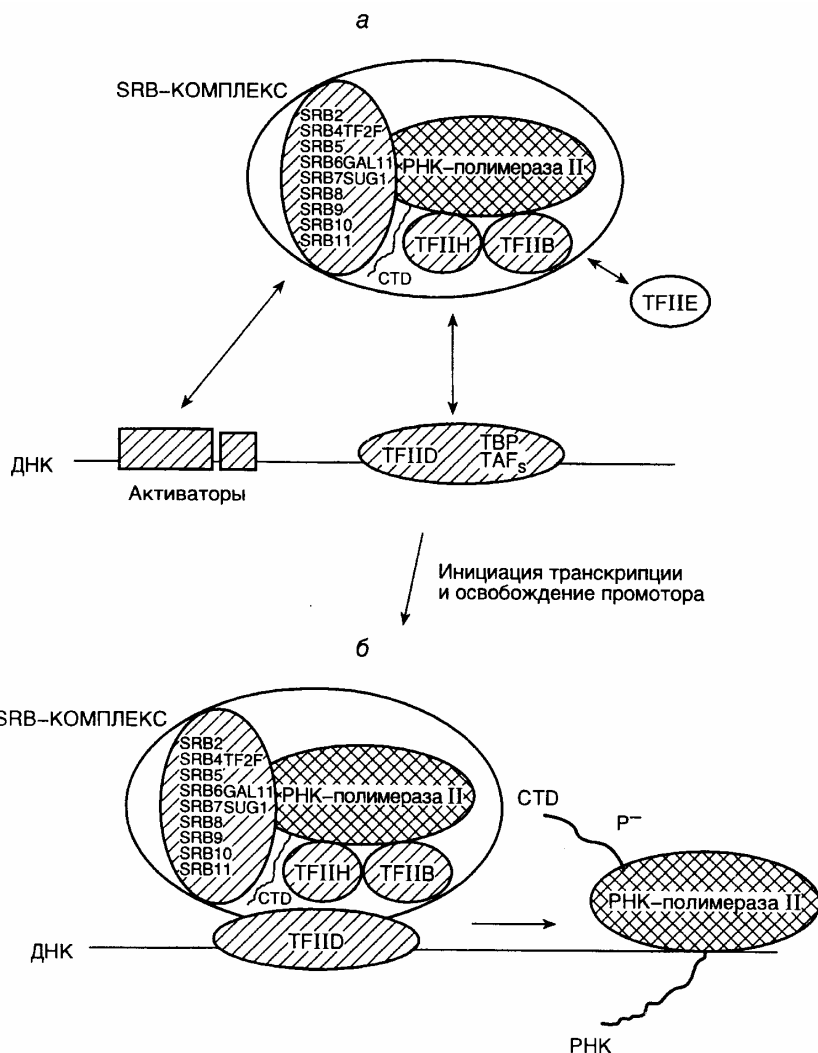


Рис. 1.7. Схема инициации транскрипции РНК-полимеразой II и освобождения промотора

а – сборка прединициационного транскрипционного комплекса; **б** – инициация транскрипции и освобождение промотора

Ко второму классу относятся факторы транскрипции, специфически взаимодействующие с определенными последовательностями ДНК, которые являются основными регуляторами транскрипции и обеспечивают тканеспецифическую экспрессию генов. И, наконец, третий класс факторов транскрипции (в том числе многочисленные TAF-белки (TAF-associated factors)) представлен недавно открытыми белками – *коактиваторами транскрипции*, которые действуют согласованно с основными и тканеспецифическими

факторами, обеспечивая более тонкую регуляцию транскрипции. О двух последних классах факторов транскрипции речь пойдет в разделе 3.2.2.

Первым с TATA-последовательностью промотора взаимодействует белковый комплекс TFIID (transcription factor II D), в состав которого входят белок, связывающий TATA-последовательность (TBP), а также еще по меньшей мере девять белковых субъединиц (см. рис. 1.7,а). Белок TBP необходим и для осуществления транскрипции РНК-полимеразами I и III, он является универсальным фактором транскрипции у эукариот. Взаимодействие TBP с TATA-последовательностью характеризуется значительным изменением структуры ДНК в этом месте, сопровождаемым частичным разворачиванием двойной спирали. После взаимодействия с TATA-последовательностью TFIID приобретает способность ассоциироваться с факторами TFIIA и TFIIB. Присутствие TFIIB делает возможным вхождение РНК-полимеразы II в прединициационный комплекс. Присоединение РНК-полимеразы II к комплексу сопровождается связыванием с ней дополнительных факторов TFIIЕ, TFIIF, TFIIH и TFIIJ, что завершается образованием так называемого *закрытого* прединициационного комплекса (см. рис. 1.7,б). Свое название комплекс получил благодаря тому, что участок ДНК промотора, входящий в его состав, в основном сохраняет свою исходную вторичную структуру – обе цепи ДНК связаны друг с другом водородными связями. В инициационном комплексе происходит локальное плавление, т.е. раскрытие соответствующего участка ДНК, и комплекс становится способным к инициации транскрипции. Аминокислотные последовательности отдельных субъединиц факторов TFIIB, TFIIЕ и TFIIF обнаруживают явную гомологию с последовательностями σ 70-фактора РНК-полимеразы *E. coli*. Это может указывать на определенную общность функций и эволюционного происхождения про- и эукариотических факторов транскрипции.

После сборки прединициационный закрытый комплекс претерпевает температурно-зависимые конформационные изменения, что необходимо для образования активного иницирующего транскрипционного комплекса. Показано, что в закрытом комплексе холофермент РНК-полимеразы *E. coli* взаимодействует с молекулой ДНК лишь с одной ее стороны, тогда как в промежуточном и *открытом* комплексах β - и β' -субъединицы РНК-полимеразы полностью охватывают молекулу ДНК в окрестностях точки инициации

транскрипции. Конформационный переход промежуточного прединициационного комплекса в открытый сопровождается локальным плавлением двойной спирали ДНК между нуклеотидами в положениях -10 и $+3$ с образованием коротких одноцепочечных участков ДНК. При наличии открытого инициационного комплекса в присутствии четырех рибонуклеозидтрифосфатов может происходить инициация транскрипции, сопровождаемая синтезом коротких (до 7–8 нуклеотидов) олигорибонуклеотидов, непрерывно освобождающихся из транскрипционного комплекса в том случае, если синтез РНК не может быть продолжен (например из-за отсутствия очередного нуклеотида). Эта стадия синтеза РНК получила название *абортивной транскрипции*. После синтеза РНК длиной более девяти нуклеотидов фермент покидает промотор, у РНК-полимеразы *E. coli* происходит отделение σ -фактора от инициационного комплекса, формируется стабильный элонгирующий комплекс, и реакция транскрипции вступает в фазу *элонгации*.

Фаза абортивной транскрипции характерна для прокариотических и эукариотических РНК-полимераз. Однако переход от стадии абортивной транскрипции к продуктивной элонгации у эукариот характеризуется рядом особенностей. В отличие от РНК-полимеразы *E. coli* (а также эукариотических РНК-полимераз I и III), РНК-полимераза II содержит в С-концевом домене большой субъединицы гексапептидный повтор (до 52 повторяющихся последовательностей) CTD (carboxy terminal domain), который является субстратом протеинкиназы. В прединициационном комплексе эта последовательность аминокислотных остатков частично фосфорилирована, тогда как у активно элонгирующей РНК-полимеразы CTD фосфорилирован полностью по остаткам Ser и Thr. Несмотря на то, что *in vitro* многие протеинкиназы обладают способностью фосфорилировать CTD, биологические функции в данном процессе приписывают протеинкиназе фактора TFIIN. Дополнительные исследования показали, что эта протеинкиназа идентична киназе *cdk7*, участвующей в регуляции клеточного цикла. Предполагается, что *in vivo* частичное фосфорилирование CTD-домена удерживает молекулу РНК-полимеразы II на промоторе, и его полное фосфорилирование необходимо для того, чтобы фермент покинул промотор и перешел к элонгации вновь синтезируемых цепей РНК. Этот процесс получил название "*освобождение промотора*" (promoter clearance (escape)) (см. рис. 1.7,б). Кроме того, переход

закрытого прединициационного комплекса, образованного РНК-полимеразой II, в открытый инициационный комплекс является АТФ-зависимым процессом.

В заключение необходимо еще раз вспомнить о том, что многие клеточные белки могут выполнять более одной функции, что является основой тесного сопряжения биохимических реакций, вовлеченных в процессы реализации генетической информации. В частности, многие (пять из девяти) субъединицы TFIIH служат основными компонентами системы эксцизионной репарации ДНК (см. табл. 1.4, а также раздел 5.2.2). Такое сопряжение приводит к более эффективной репарации повреждений ДНК в активно транскрибируемых генах по сравнению с молчащими последовательностями ДНК.

Элонгация цепей РНК. Момент перехода РНК-полимеразы от инициации транскрипции к элонгации точно не определен. Три основных биохимических события характеризуют этот переход в случае РНК-полимеразы *E. coli*: отделение σ -фактора, первая транслокация молекулы фермента вдоль матрицы и сильная стабилизация транскрипционного комплекса, который кроме РНК-полимеразы включает растущую цепь РНК и транскрибируемую ДНК. Все те же явления характерны и для РНК-полимераз эукариот, хотя в этом случае момент перехода от инициации к элонгации еще более неопределен. И в том, и в другом случае переход от инициации к элонгации сопровождается разрывом связей между ферментом, промотором, факторами инициации транскрипции, а в ряде случаев – переходом РНК-полимеразы в состояние компетентности в отношении элонгации (например фосфорилирование CTD-домена у РНК-полимеразы II). Фаза элонгации заканчивается после освобождения растущего транскрипта и диссоциации фермента от матрицы (терминация).

Биохимические особенности элонгации. РНК-полимеразы являются процессивными ферментами: если фермент освобождает растущую цепь РНК до завершения транскрипции гена, то он не в состоянии связать эту РНК вновь и продолжить транскрипцию. Поэтому каждый новый раунд синтеза РНК начинается с реинициации транскрипции на промоторе. Такие биохимические особенности РНК-полимераз объясняют, почему стабильность транскрибирующего комплекса является одним из его критических свойств.

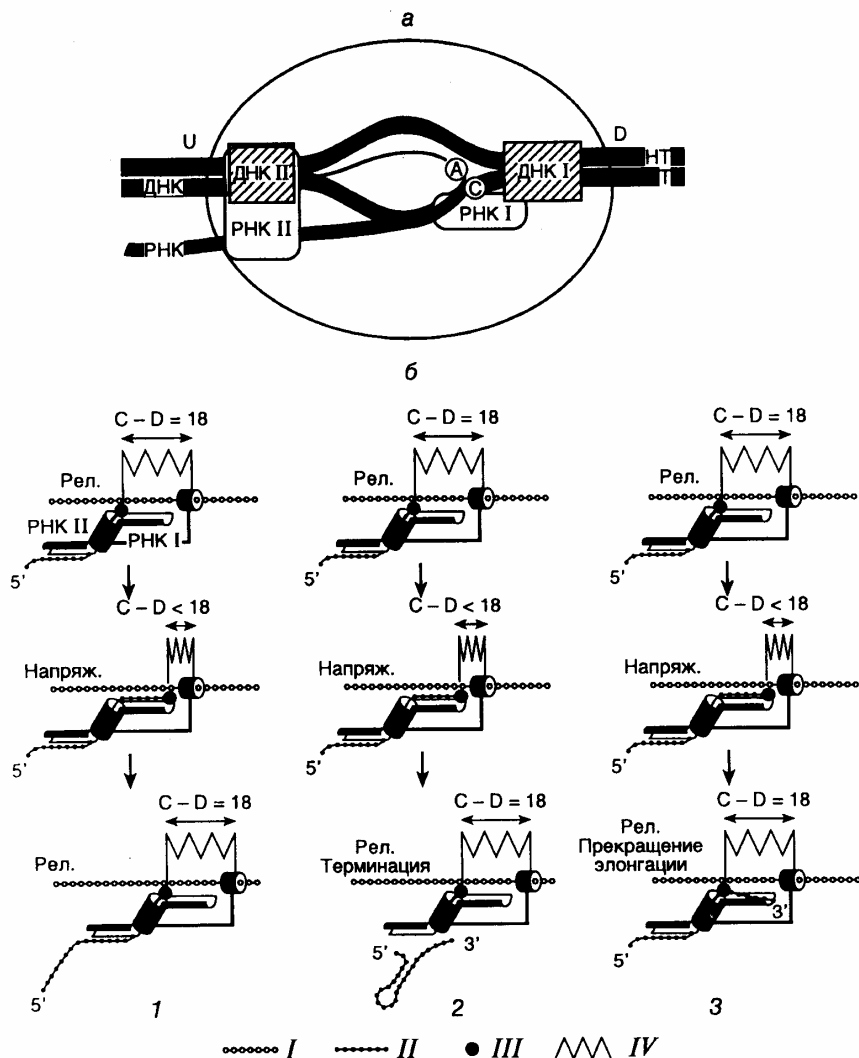


Рис. 1.8. Модели структуры тройного комплекса элонгирующей РНК-полимеразы *E. coli* (а) и механизма скачкообразной элонгации (б)
а: Т и НТ – соответственно транскрибируемая и нетранскрибируемая цепи ДНК. ДНК I и ДНК II, РНК I и РНК II – сайты связывания соответственно ДНК и РНК, А – активный центр, D и U – соответственно передняя и задняя части движущегося фермента, С – 3'-конец РНК.
б: Цикл скачкообразной элонгации транскрипции (1), сопровождаемой терминацией синтеза РНК (2) или образованием комплексов полностью прекративших элонгацию (3). I и II – соответственно транскрибируемая цепь ДНК и строящаяся цепь РНК; III – активный центр фермента; IV – напряженное (напряж.) и релаксированное (рел.) состояния полипептидной цепи РНК-полимеразы; цилиндр, охватывающий цепь ДНК

– ДНК-связывающий центр I; C–D – расстояние между 3'-концом РНК и передней границей РНК-полимеразы (п.о.)

Действительно, многие гены эукариот обладают очень большими размерами (длина гена дистрофина человека, например превышает 2000 т.п.о. и его транскрипция продолжается ~ 17 ч), следовательно, преждевременная терминация транскрипции могла бы с большой вероятностью прекратить экспрессию таких генов вообще. Из этих соображений не удивительно, что тройной комплекс РНК-полимераза–РНК–матрица чрезвычайно стабилен. Такие комплексы устойчивы к высокой ионной силе, допускают очистку гель-фильтрацией, преципитацию антителами, остаются активными после проведения электрофореза в неденатурирующих условиях и могут храниться в отсутствие рибонуклеозидтрифосфатов при +4°C в течение нескольких дней без потери активности.

В настоящее время разработаны методы, позволяющие задерживать элонгирующую РНК-полимеразу бактерий в любом участке матрицы путем удаления из реакционной смеси одного из рибонуклеозидтрифосфатов. При таком подходе РНК-полимеразу иммобилизуют на носителе, и процесс транскрипции осуществляют непосредственно в колонке, последовательно добавляя нуклеотид за нуклеотидом после отмывания предыдущего. Это дает возможность последовательно перемещать РНК-полимеразу от одного нуклеотида к другому вдоль матрицы, как бы шагать вдоль матричной ДНК (walking). Такой метод является прекрасным инструментом для исследования механизмов элонгации и обнаружения факторов транскрипции, оказывающих влияние на этот процесс.

Скорости транскрипции матричной ДНК РНК-полимеразами сильно различаются. Обнаружена слабая корреляция между субъединичным строением РНК-полимераз и скоростью, с которой они способны элонгировать ДНК. Так, РНК-полимеразы бактериофагов, состоящие из одной субъединицы, являются наиболее быстрыми среди ДНК-зависимых РНК-полимераз. Они способны элонгировать *in vitro* растущую цепь РНК со скоростью 200–400 нт/с. Бактериальные РНК-полимеразы транскрибируют ДНК с промежуточной скоростью – 50–100 нт/с, тогда как скорость элонгации РНК-полимеразой II многоклеточных организмов *in vitro* составляет всего 5–10 нт/с. Эукариотические РНК-полимеразы элонгируют цепи РНК *in vivo* со скоростью

~20–30 нт/с.

Точность транскрипции, осуществляемой РНК-полимеразами, ниже соответствующего параметра ДНК-полимераз. Частота ошибок в виде включенных в РНК некомплементарных матрице нуклеотидов составляет 10^{-3} – 10^{-5} . Это объясняется, прежде всего, тем, что РНК-полимеразы лишены корректирующей системы в виде 3'→5'- и 5'→3'-экзонуклеазной активности. В определенных условиях РНК-полимеразы обладают способностью отщеплять 3'-концевую часть транскрипта, находящегося в составе тройного комплекса (см. ниже). Однако до сих пор не ясно, имеет ли эта активность отношение к коррекции ошибок. С эволюционной точки зрения понятно, почему точность транскрипции в меньшей степени подвержена давлению отбора по сравнению с точностью репликации. Действительно, если ошибки репликации должны сопровождаться изменением генотипа клетки или организма, то ошибки транскрипции не приводят к таким последствиям.

Структура элонгирующего комплекса. Основные черты структуры тройного комплекса, осуществляющего элонгацию цепей РНК, консервативны у всех ДНК-зависимых РНК-полимераз (см. рис. 1.8). В каждом элонгирующем комплексе имеются каталитический центр, одноцепочечная область ДНК-матрицы, а также несколько сайтов связывания ДНК и РНК. Для реализации принципа комплементарности при построении растущей цепи РНК участок матричной ДНК, входящий в состав комплекса, находится в расплавленном состоянии, и одна из его цепей служит матрицей при транскрипции. Этот участок ДНК, называемый *транскрипционным пузырьком*, или *транскрипционной сферой* (transcription bubble), контактирует с каталитическим центром РНК-полимеразы. По обеим сторонам транскрипционного пузырька имеются участки ДНК, которые при перемещении фермента вдоль матрицы подвергаются плавлению (расплетанию) и повторному отжигу, в результате которого восстанавливается исходная структура ДНК. Считается, что этот процесс не является каталитическим, протекает самопроизвольно и связан с особенностями структуры РНК-полимеразы как таковой.

В молекуле РНК-полимеразы имеются несколько участков связывания матричной ДНК и участок связывания растущей цепи РНК, которые обеспечивают точное расположение 3'-ОН-конца РНК относительно

транскрибируемого в данный момент нуклеотида матрицы и соответствующего субстрата. В ранних моделях элонгирующего комплекса постулировалось существование короткого гибрида между матричной ДНК и строящейся цепью РНК, однако в настоящее время эта модель подвергается сомнению. Считается, что если гибрид существует, то его длина составляет лишь 2–12 п.о. Современные модели элонгирующего комплекса предполагают наличие механизма, обеспечивающего выбор правильного нуклеотида из пула субстратов. Хотя в принципе этот процесс может быть исключительно следствием комплементарного взаимодействия нуклеотида и матрицы, более распространена точка зрения, согласно которой ключевую роль здесь играет сама молекула РНК-полимеразы, пространственная структура которой формируется при ее контакте с ДНК и растущим транскриптом.

В современных моделях структуры элонгирующего комплекса предполагают наличие в РНК-полимеразе двух сайтов, удерживающих растущую цепь РНК, и двух участков связывания ДНК-матрицы. В совокупности эти участки связывания обеспечивают замечательную стабильность комплекса. Тем не менее, точное положение этих сайтов связывания нуклеиновых кислот в настоящее время неизвестно. Для РНК-полимеразы *E. coli* имеются доказательства того, что задняя часть элонгирующего фермента контактирует с двухцепочечной ДНК, а передняя – с одноцепочечной, причем с ее транскрибируемой (матричной) цепью.

Способы блокировки элонгации растущих цепей РНК. Элонгация РНК не происходит с постоянной скоростью. Во время элонгации РНК-полимераза может получать сигналы, вызывающие *задержку* транскрипции (pause), ее *прекращение* (arrest), или терминацию. В результате задержки молекула РНК-полимеразы временно останавливает синтез РНК на определенный период времени, после чего спонтанно продолжает транскрипцию. В отличие от этого, элонгирующие комплексы, прекратившие синтез РНК, неспособны продолжить транскрипцию без помощи дополнительных факторов. Сигналы, вызывающие такое поведение РНК-полимеразы, могут быть внутренними, в виде определенной последовательности нуклеотидов транскрибируемой ДНК, или внешними – в виде специфических белков-регуляторов, связавшихся с матрицей.

Прекращение транскрипции может быть вызвано и в искусственных условиях при отсутствии одного или нескольких рибонуклеозидтрифосфатов. Как во время задержки, так и во время прекращения синтеза РНК РНК-полимераза остается каталитически активной, стабильно связанной с ДНК и удерживает растущий транскрипт в составе тройного комплекса. Именно эти характерные особенности отличают РНК-полимеразу в состоянии задержки или прекращения синтеза РНК от фермента, находящегося в фазе терминации транскрипции. Во многих случаях, когда невозможно четко дифференцировать три вышеупомянутых состояния транскрипции, говорят просто о блокировке синтеза РНК.

Как остановка, так и прекращение синтеза РНК характеризуются двумя параметрами – временем полужизни заблокированного состояния комплекса и эффективностью распознавания сигнала. Под временем полужизни блокировки элонгации понимают промежуток времени, за который половина заблокированных молекул РНК-полимеразы вступает в фазу продолжения синтеза РНК. Остановка может быть настолько кратковременной, что время полужизни заблокированного комплекса невозможно точно определить; в то же время задержка элонгации может продолжаться в течение нескольких минут, что зависит от локализации и структуры соответствующего сайта на матричной ДНК. Эффективность распознавания сигнала блокировки определяется долей молекул РНК-полимеразы, отвечающих прекращением элонгации на соответствующем сайте, и может принимать значения от нескольких до 90%. Описано несколько сайтов, обладающих эффективностью 100%.

Временные остановки элонгирующей РНК-полимеразы на транскрибируемой ДНК бактериофага Т7 были обнаружены в бесклеточных системах, синхронизированных в отношении инициации синтеза РНК. Неожиданно оказалось, что молекулы РНК-полимеразы, синхронно иницировавшие транскрипцию с одного промотора, быстро и специфически распределяются в составе тройного комплекса вдоль матрицы, оставаясь ассоциированными с растущими цепями РНК дискретной длины. Исследование этого феномена привело к заключению о временных остановках РНК-полимеразы во время транскрипции в определенных участках матрицы и было одним из первых доказательств того, что синтез РНК не является непрерывным процессом. Такие временные задержки во время элонгации характерны для

многих РНК-полимераз, включая РНК-полимеразы I, II и III эукариотических организмов. Как у прокариот, так и у эукариот имеются многочисленные механизмы, регулирующие эффективность задержки транскрипции в определенных сайтах транскрибируемой ДНК (подробнее см. раздел 3.2).

В отличие от описанной выше временной задержки транскрипции, полное прекращение элонгации РНК внутри транскрибируемого гена физически блокирует перемещение вдоль гена других молекул РНК-полимеразы, инициировавших транскрипцию на тех же промоторах. Следствием этого является полное прекращение транскрипции такого гена, что служит одним из распространенных механизмов регуляции экспрессии генов на уровне элонгации РНК. Состояние прекращения элонгации РНК-полимеразой может возникать на некоторых участках матрицы *in vitro* в отсутствие одного или нескольких рибонуклеозидтрифосфатов. В этом случае, в отличие от временной задержки транскрипции, добавление в бесклеточную систему недостающих нуклеотидов не приводит к возобновлению синтеза РНК. Большинство тройных комплексов, прекративших транскрипцию по такому механизму, являются стабильными, а входящая в их состав РНК-полимераза полностью сохраняет свою активность. Переход элонгирующих комплексов из состояния временной задержки в состояние полного прекращения транскрипции происходит постепенно, по мере увеличения времени инкубации бесклеточной системы в отсутствие субстратов. В настоящее время не существует прямых доказательств наличия феномена полного прекращения элонгации *in vivo*. Однако по ряду косвенных данных, включая открытие белков, реактивирующих такие тройные комплексы, считается, что полное прекращение транскрипции имеет место также в клетках живого организма и является одним из регуляторных механизмов, координирующих экспрессию генов на уровне транскрипции. Задержка и полное прекращение транскрипции во время элонгации цепей РНК становятся возможными благодаря взаимодействиям между РНК-полимеразой, растущей цепью РНК и транскрибируемой ДНК и не требуют участия дополнительных факторов. Несмотря на интенсивные исследования, до сих пор не удается предсказывать места задержки элонгации РНК на матричных ДНК.

Природа сигналов, обеспечивающих задержку элонгирующих

тройных комплексов у прокариот. Сигналы, вызывающие задержку элонгации, многокомпонентны. Они включают элементы как РНК, так и ДНК. Во многих сайтах задержки элонгации РНК-полимераза синтезирует участок РНК, который имеет склонность образовывать стабильную вторичную структуру. В этом случае мутации, разрушающие шпильку, уменьшают время полужизни комплекса в состоянии задержки элонгации. И наоборот, мутации, восстанавливающие вторичную структуру, увеличивают время полужизни таких комплексов. Это указывает на то, что сама вторичная структура, а не конкретная последовательность нуклеотидов внутри шпильки существенна для задержки элонгации растущей цепи РНК. В пользу этого вывода свидетельствует и тот факт, что включение в РНК IMP вместо GMP, приводящее к снижению числа водородных связей в шпильках, а следовательно, и их прочности, приводит к элиминации некоторых мест задержки элонгации. Тем не менее, образование шпильки в РНК еще не является достаточным условием задержки транскрипции в соответствующем участке матрицы, так как многие места задержки не содержат таких последовательностей. Последовательности ДНК и РНК, расположенные между основанием шпильки и активным центром РНК-полимеразы, также оказывают влияние на задержку элонгации. Замены нуклеотидов в 3'-концевых частях таких последовательностей сопровождаются изменениями времени полужизни комплекса в состоянии задержки. Полагают, что в этом случае вторичная структура РНК не имеет отношения к наблюдаемым эффектам.

Сигналы задержки элонгации могут действовать не только через РНК, но и на уровне ДНК. Так, последовательности, расположенные даже на 17 нуклеотидов ниже участка задержки транскрипции, оказывают влияние на задержку. Полагают, что одной из причин этого может быть воздействие таких последовательностей на способность РНК-полимеразы создавать локально расплетенный участок матричной ДНК. Альтернативным объяснением является влияние таких участков матричной ДНК на конформацию РНК-полимеразы. В случае РНК-полимеразы *E. coli* эти участки ДНК находятся в контакте с ферментом в одном из его ДНК-связывающих сайтов. С помощью мутационного анализа было установлено, что на задержку элонгации могут оказывать влияние и последовательности нуклеотидов нематричной цепи ДНК.

Природа прокариотических сигналов, вызывающих прекращение

элонгации РНК. Так же как и в предыдущем случае, природа сигналов, вызывающих прекращение элонгации, не вполне ясна. В окрестностях сайтов на ДНК, вызывающих прекращение элонгации, не обнаружено каких-либо канонических последовательностей. К этому феномену может быть причастна вторичная структура РНК, поскольку замещение GMP на IMP, приводящее к ослаблению вторичной структуры РНК, во многих случаях предотвращает прекращение транскрипции в соответствующих участках ДНК. Большинство элонгирующих комплексов, прекративших транскрипцию, было получено в искусственных условиях при недостатке рибонуклеозидтрифосфатов. В отличие от сигналов задержки транскрипции, ни один из сигналов прекращения транскрипции не был обнаружен внутри прокариотических генов.

Сигналы остановки транскрипции на уровне элонгации у эукариот. Участки ДНК, вызывающие задержку и прекращение элонгации у эукариот, охарактеризованы слабо. Имеются данные, что в случае РНК-полимераз I и II такие участки ДНК часто содержат Т-богатые последовательности в нематричной цепи, однако это не является общим правилом. Если такие Т-богатые участки существуют, то в них часто происходит формирование изгиба ДНК, а эффективность блока прямо пропорциональна длине Т-богатой последовательности. Наличие шпилек в транскриптах этих участков не всегда требуется для блокировки транскрипции. В случае РНК-полимеразы II существенным моментом блокировки является взаимодействие каталитического сайта фермента с ниже расположенными последовательностями ДНК. При этом задержанный комплекс с течением времени может превращаться в комплекс, полностью прекративший транскрипцию. Эффективность распознавания РНК-полимеразой участков задержки на ДНК зависит от свойств промотора, на котором произошла инициация транскрипции, а также расстояния сайта блокировки от промотора. Как у прокариот, так и у эукариот специфические белки оказывают сильное влияние на способность РНК-полимеразы распознавать участки задержки и прекращения транскрипции. Многие из этих белков участвуют в регуляции транскрипции на уровне элонгации.

Внешние факторы, блокирующие элонгацию цепей РНК. В дополнение к только что рассмотренным внутренним сигналам, препятствующим элонгации РНК, существует большое число внешних

факторов, оказывающих влияние на этот процесс. Имеется много примеров, когда ДНК-связывающие белки блокируют элонгацию цепей РНК вирусными, прокариотическими и эукариотическими РНК-полимеразами как в природных условиях, так и в бесклеточных системах. Однако во многих случаях РНК-полимеразы способны обходить ДНК-связывающие белки без задержки транскрипции. Хотя белки являются основным препятствием на пути элонгации РНК, имеются и другие факторы, оказывающие подобное влияние на элонгацию. В частности, показано, что интенсивное негативное и позитивное суперскручивание ДНК-матрицы может увеличивать время полужизни элонгирующих комплексов в задержанном состоянии. ДНК в Z-форме также может ингибировать элонгацию. Многие вещества, взаимодействующие с ДНК, а также повреждения ДНК ингибируют элонгацию. Интересно, что РНК-полимераза *E. coli* способна преодолевать участки ДНК с отсутствующими азотистыми основаниями. Еще более удивительной является способность T7-РНК-полимеразы не замечать короткие бреши в матричной цепи ДНК длиной в 1–5 нуклеотидов. Она может перебираться через бреши даже длиной в 24 нуклеотида, хотя и с меньшей эффективностью. Прохождение РНК-полимеразы через такие бреши сопровождается образованием единого транскрипта, содержащего делецию в соответствующем участке РНК. Способность РНК-полимеразы транскрибировать ДНК, содержащие бреши, зависит от природы 3'-концевого нуклеотида, после которого начинается брешь: 3'-ОН-конец допускает транскрипцию, тогда как 3'-фосфат блокирует элонгацию.

Расщепление транскрипта в комплексах, полностью прекративших элонгацию. РНК-полимеразы в составе тройных комплексов, включая все три РНК-полимеразы эукариот и РНК-полимеразу *E. coli*, способны осуществлять эндонуклеазное расщепление растущей цепи РНК вблизи ее 3'-конца. Образующийся в результате более протяженный 5'-концевой фрагмент РНК остается прочно связанным с ферментом, тогда как 3'-концевой фрагмент освобождается из тройного комплекса. При этом каталитический центр РНК-полимеразы остается правильно ориентированным в отношении оставшегося транскрипта и ДНК-матрицы, и она способна продолжать транскрипцию, не образуя в синтезируемой РНК внутренних делеций. Реакция расщепления не является пирофосфоролизом, поскольку в результате пирофосфоролиза формируется фрагмент РНК, содержащий 5'-концевой нуклеозидтрифосфат, а

при расщеплении транскрипта РНК-полимеразой образуется фрагмент РНК с монофосфатом на 5'-конце. Белки *E. coli* GreA и GreB, а также эукариотический фактор транскрипции TFIIIS стимулируют процесс расщепления. Те же факторы и, кроме того, состав тройного комплекса оказывают влияние на размер освобождающегося в процессе расщепления 3'-концевого фрагмента РНК, который может варьировать от 1 до 17 нуклеотидов. В разных тройных комплексах процесс расщепления может происходить мгновенно или растягиваться на несколько часов. Полагают, что кинетика этого процесса зависит от конформации тройного комплекса, которая может быть различной в разных участках матричной ДНК.

Имеются доказательства того, что расщепление транскрипта катализируется одним активным центром фермента, ответственным, кроме того, за полимеризацию рибонуклеозидтрифосфатов и пирофосфоролиз РНК. В частности, было показано, что неорганический пирофосфат может индуцировать освобождение олигонуклеотидного фрагмента РНК, содержащего 5'-концевую трифосфатную группу, из тройного комплекса, прекратившего синтез РНК. Эти данные интерпретируются в пользу того, что активный центр РНК-полимеразы, прекратившей транскрипцию, смещается от 3'-конца элонгируемого транскрипта к его 5'-концу, где он и катализирует пирофосфоролитическое расщепление РНК. Кроме того, до настоящего времени не обнаружено ингибиторов расщепления транскрипта, которые одновременно не подавляли бы самого синтеза РНК. Не найдено и мутантов, у которых оба этих процесса были бы разделены. И, наконец, расщепление транскрипта и полимеризация субстратов происходят в одинаковых условиях: в присутствии двухвалентных катионов с одним и тем же оптимумом рН реакционной смеси.

Физиологическое значение реакции расщепления транскрипта РНК-полимеразой еще предстоит выяснить. Известно, что факторы GreA и TFIIIS, которые стимулируют расщепление РНК, одновременно подавляют образование прекративших транскрипцию комплексов. В соответствии с наиболее распространенной точкой зрения расщепление транскрипта помогает РНК-полимеразе выйти из состояния полного прекращения транскрипции. Полагают, что в тройном комплексе, прекратившем транскрипцию, активный центр РНК-полимеразы постепенно смещается от 3'-конца РНК к 5'-концу и

утрачивает способность ее элонгировать. Эндонуклеазное расщепление транскрипта правильно ориентирует новый 3'-концевой нуклеотид РНК относительно активного центра РНК-полимеразы, что реактивирует процесс транскрипции. Другой возможной функцией расщепления РНК молекулой РНК-полимеразы может быть коррекция ошибочно включенных в растущую РНК 3'-концевых нуклеотидов. Если после включения некомплементарного матрице нуклеотида в РНК молекула РНК-полимеразы делает паузу в элонгации, то вырезание 3'-концевого фрагмента растущей цепи РНК и продолжение элонгации можно рассматривать как имитацию РНК-полимеразой хорошо известной 3'→5'-экзонуклеазной корректирующей активности ДНК-полимеразы I *E. coli*. Такой механизм может помогать РНК-полимеразе в преодолении физических препятствий во время транслокации. Известно, что места задержки транскрибирующей РНК-полимеразы на ДНК распознаются молекулами фермента, находящимися в определенной фазе (рамке считывания) по отношению к этим сайтам. Поскольку отщепление 3'-концевого участка ДНК и обратная транслокация могут перевести фермент в новую фазу по отношению к сайтам задержки на ДНК, то фермент, начиная транскрипцию из нового положения, сможет беспрепятственно преодолевать сайты задержки в процессе элонгации. Одним из таких преодолеваемых при транскрипции барьеров для РНК-полимеразы у эукариот могут быть, в частности нуклеосомы.

Помимо всего прочего, в настоящее время показано, что вышеупомянутые факторы GreA и GreB стимулируют процесс освобождения промоторов бактериальными РНК-полимеразами.

Реитеративный синтез гомополимеров. Молекулы РНК, обычно синтезируемые РНК-полимеразами, строго комплементарны транскрибируемому участку матричной РНК. Однако в редких случаях *in vivo*, а также в определенных искусственных условиях РНК-полимеразы способны осуществлять так называемый *реитеративный синтез гомополимеров*, во время которого молекула РНК-полимеразы, многократно проскальзывая вдоль короткой гомополимерной последовательности (transcriptional slippage), синтезирует длинные комплементарные этой последовательности гомополимеры – поли(А), поли(У) или поли(С). Размеры образующихся при этом гомополимерных продуктов могут более чем в 10–20 раз превышать длину

матричной последовательности. Реитеративный синтез протекает особенно эффективно на одноцепочечных ДНК в присутствии ионов Mn^{2+} вместо Mg^{2+} и необходимого рибонуклеозидтрифосфата. Если же олигонуклеотидная последовательность матрицы включена в случайную последовательность, то реитеративного синтеза не происходит при наличии в реакционной смеси кроме нуклеотида, составляющего гомополимер, других рибонуклеозидтрифосфатов. Это обусловлено пребыванием тройного комплекса в состоянии обычной элонгации.

Реитеративный синтез гомополимеров РНК-полимеразами может происходить как во время инициации транскрипции, так и в фазе элонгации. При инициации РНК-полимераза может начинать реитеративный синтез при наличии в окрестностях точки инициации транскрипции гомополимера длиной в 3 и даже 2 нуклеотида. Во время элонгации РНК-полимераза предпочитает осуществлять реитеративный синтез на более длинных гомополимерных последовательностях. В частности, в двух исследованных случаях элонгирующая РНК-полимераза *E. coli* осуществляла реитеративный синтез поли(А) или поли(У) на последовательностях, состоящих из dТ или dА, длиной в 10 нуклеотидов.

Несмотря на то что реитеративный синтез гомополимеров является редким событием *in vivo*, он выполняет определенные функции в регуляции транскрипции. Показано, что проскальзывание РНК-полимеразы вдоль гомополимерных последовательностей матрицы может быть механизмом, изменяющим рамки считывания мРНК при трансляции, редактирующим генетическую информацию мРНК, осуществляющим присоединение к РНК поли(А)-хвостов и 5'-концевых последовательностей к вирусным РНК, а также регулирующим инициацию транскрипции некоторых бактериальных оперонов (например *pyrBI* или *codBA*). Поскольку нерегулируемый реитеративный синтез внутри кодирующих частей генов таит большую опасность для их экспрессии, предполагают, что гомополимерные последовательности в транскрипционных единицах являются мишенью для эволюционного отбора, который их элиминирует. Действительно, компьютерный анализ под этим углом зрения последовательностей нуклеотидов *E. coli* показал отсутствие достаточно протяженных гомополимеров в кодирующих частях генов и их наличие в некодирующих последовательностях бактериального генома.

Современные модели структуры элонгирующих тройных комплексов и механизма элонгации растущих цепей РНК. В 1992 г. М. Чамберлин с сотрудниками разработали модель элонгации РНК, которая содержала целый ряд новых идей. Прежде всего было постулировано, что процессы транслокации РНК-полимеразы вдоль ДНК и присоединение нуклеотидов к растущей цепи РНК в активном центре фермента разделены во времени. Это разделение становится возможным благодаря наличию в молекуле РНК-полимераз двух ДНК-связывающих сайтов, каждый из которых перекрывает на матричной ДНК ~ 10 нуклеотидных пар (см. рис. 1.8). Сайт I локализован в задней по отношению к направлению транскрипции части фермента, а сайт II – в передней. Предполагалось также, что эти два сайта могут перемещаться вдоль ДНК в процессе элонгации независимо друг от друга. В соответствии с предложенной моделью, молекула РНК-полимеразы перемещается вдоль ДНК наподобие гусеницы: когда один из сайтов связывания ДНК фиксирован, другой может двигаться вперед. Еще одной новой чертой модели было предположение о том, что растущая цепь РНК удерживается внутри элонгирующего комплекса не 12-нуклеотидным ДНК–РНК-гибридом, а двумя сайтами связывания самой РНК-полимеразы, каждый из которых перекрывает ~10 нуклеотидов матрицы. При этом длина участка РНК, заключенного между сайтами связывания, составляет 30–40 нуклеотидов. Предполагалось также, что перемещение каталитического центра молекулы РНК-полимеразы сопряжено с движением переднего сайта связывания ДНК II и может происходить независимо от переноса заднего сайта I вперед вдоль РНК-связывающего сайта I во время добавления нуклеотидов к растущей цепи РНК, сопровождаемого заполнением этого сайта (см. рис. 1.8,б,1).

Согласно предложенной модели элонгация цепей РНК представляется в виде циклического процесса. В начале цикла каталитический центр молекулы РНК-полимеразы располагается у задней границы РНК-связывающего сайта I в соответствии с положением 3'-ОН-конца РНК. Последовательно присоединяя нуклеотиды к растущей цепи РНК, каталитический участок перемещается относительно РНК-связывающего сайта I и в конце концов заполняет этот сайт десятью нуклеотидами вновь синтезированного участка РНК (см. рис. 1.8,б,1). Во время этой фазы элонгации ДНК-связывающий сайт I остается фиксированным на ДНК, тогда как ДНК-связывающий сайт II перемещается

вперед синхронно с каталитическим участком на десять нуклеотидов. В конце фазы добавления нуклеотидов ДНК- и РНК-связывающие сайты II фиксируются на своих лигандах, а ДНК-связывающий сайт I переносится вперед на десять нуклеотидов в новое фиксированное положение. Это перемещение освобождает РНК-связывающий сайт I, делая его готовым к повторению цикла транслокации.

Волна исследований, последовавшая за появлением этих новых идей, быстро подтвердила реальность основных положений модели элонгации транскриптов, выдвинутых группой М. Чамберлина, и дополнила ее рядом существенных моментов. Оказалось, что перемещение тройного комплекса вдоль транскрибируемой ДНК не всегда скачкообразно. Большую часть транскрибируемых ДНК молекулы РНК-полимеразы проходят монотонно, регулярно присоединяя к растущим цепям РНК нуклеотид за нуклеотидом. Прерывистая, скачкообразная элонгация РНК имеет место лишь на участках матричной ДНК, в которых происходит задержка транскрипции или ее полное прекращение. Встретив препятствие на своем пути в виде специфической последовательности нуклеотидов или белков, ассоциированных с ДНК, молекула РНК-полимеразы в составе тройного комплекса переходит во внутренне напряженное состояние, которое она может разрешить тремя путями: или преодолеть препятствие, переместив свою переднюю границу вперед за пределы препятствия, и продолжить монотонную элонгацию транскрипта, или терминировать транскрипцию (см. рис. 1.8,б,2), или полностью прекратить синтез РНК. В последнем случае каталитический участок фермента смещается назад по отношению к синтезированной РНК и ее 3'-концевой нуклеотид становится недоступным для элонгации (см. рис. 1.8,б,3). Эндонуклеазное отщепление 3'-концевого фрагмента РНК в комплексе, прекратившем элонгацию, реактивирует транскрипцию. Очищенные РНК-полимеразы, особенно РНК-полимераза II эукариот, из-за частых задержек на матрице в процессе элонгации транскрибируют молекулы ДНК со скоростью, по крайней мере, в 10 раз меньшей, чем *in vivo*. В этом случае стимулирующее действие на элонгацию оказывают специфические факторы элонгации.

Основные факторы элонгации РНК-полимеразы II. У эукариот имеются белковые факторы элонгации РНК, которые относят к двум различным

классам: основным (general) и регуляторным факторам. Как и в случае основных факторов инициации транскрипции, основные факторы элонгации РНК обеспечивают эффективную транскрипцию всех генов, кодирующих белки, тогда как регуляторные факторы специфически контролируют экспрессию отдельных генов или даже целых их семейств. Основные факторы элонгации РНК-полимеразы II и некоторые их свойства перечислены в табл. I.5. Первыми в этой таблице упомянуты белки P-TEFb (positive transcript elongation factor b) и SII, которые препятствуют переходу задержанных элонгирующих комплексов РНК-полимеразы II в состояние полного прекращения транскрипции. Активность фактора P-TEFb, впервые

Таблица I.5

Основные факторы элонгации РНК-полимеразы II

Фактор	Структура	Молекулярная масса полипептидов, кДа	Функция
P-TEFb	Гетеродимер	124, 43	Препятствует прекращению элонгации, ингибируется DRB
SII (TFIIS)	Мономер	38	Препятствует прекращению элонгации, стимулирует расщепление транскрипта
TFIIF	Гетеродимер	30, 70	Устраняет задержку элонгации РНК
Элонгин (SIII)	Гетеротример, включающий элонгины А, В и С		Та же
Элонгин А		110	Активная субъединица
Элонгин В		18	Регуляторная субъединица
Элонгин С		15	Та же
ELL		80	Устраняет задержку элонгации РНК

Примечание. DRB – 5,6-дихлоро-1-β-D-рибофуранозилбензимидазол

выделенного в гомогенном состоянии из экстрактов *Drosophila*, подавляется нуклеотидным аналогом – DRB. Фенотипическим проявлением действия этого ингибитора является общее подавление синтеза гРНК в ядрах вследствие резкого повышения частоты перехода элонгирующего комплекса в состояние полного прекращения транскрипции вблизи промоторов. DRB не подавляет элонгацию цепей РНК в бесклеточных системах транскрипции, реконструированных из высокоочищенных компонентов, что дало основание предполагать наличие дополнительных факторов, которые контролируют процесс перехода комплексов РНК-полимеразы II в фазу элонгации и чувствительны к действию этого ингибитора. Фактор P-TEFb оказался белком, обладающим именно такими свойствами. Механизм действия P-TEFb, благодаря которому этот фактор препятствует прекращению транскрипции РНК-полимеразой II, неизвестен. Предполагают, что он может быть связан с фосфорилированием РНК-полимеразы II или сопутствующих факторов транскрипции.

Небольшой белковый фактор SII, впервые выделенный из клеток асцитной опухоли Эрлиха, обеспечивает преодоление РНК-полимеразой II препятствий в виде нуклеопротеиновых комплексов или специфических последовательностей ДНК, вызывающих преждевременное прекращение транскрипции в кодирующих частях генов. Однако он не оказывает влияния на РНК-полимеразу, прекратившую элонгацию в DRB-чувствительной фазе. Фактор SII стимулирует эндонуклеазное отщепление 3'-концевой части транскрипта в комплексе, прекратившем элонгацию, что дает возможность РНК-полимеразе II продолжить элонгацию цепи РНК. Активный сайт РНК-полимеразы, обладающий такой эндонуклеазной активностью, ингибируется α -аманитином – специфическим ингибитором РНК-полимеразы II эукариот. Клетки дрожжей, у которых фактор SII инактивирован под действием мутаций, обладают повышенной чувствительностью к 6-азаурацилу и микофеноловой кислоте, которые, как известно, ингибируют биосинтез нуклеотидов, понижая внутриклеточное содержание GTP и UTP. Это, в свою очередь, оказывает сильное влияние на эффективность элонгации РНК РНК-полимеразой II в мутантных клетках.

Другая группа основных факторов элонгации супрессирует задержку элонгации цепей РНК, тем самым уменьшая вероятность перехода

элонгирующих комплексов в состояние полного прекращения элонгации. Эту группу составляют три структурно неродственных белка: факторы TFIIIF, элонгин (SIII) и ELL, которые, по-видимому, взаимодействуют непосредственно с компонентами тройного элонгирующего комплекса. Ни один из этих белков не способен реактивировать комплексы, полностью прекратившие транскрипцию, или стимулировать расщепление РНК в этих комплексах. Точный механизм супрессирующего действия данных факторов на задержку элонгации неизвестен. Недавно было установлено, что и элонгин, и фактор TFIIIF резко повышают способность РНК-полимеразы II осуществлять зависимое от матрицы присоединение рибонуклеозидтрифосфатов к 3'-ОН-концам фрагментов ДНК, которые в этом случае выполняют функцию праймеров. Предполагают, что роль элонгина и фактора TFIIIF может заключаться в обеспечении правильного расположения в активном центре элонгирующего фермента 3'-ОН-концов растущих транскриптов. Фактор TFIIIF занимает особое место среди других основных факторов транскрипции, поскольку только он обладает способностью контролировать активность РНК-полимеразы II как на стадии инициации транскрипции, так и в фазе элонгации. При этом способность этого фактора оказывать действие на инициацию транскрипции или элонгацию контролируется разными доменами его полипептидных цепей.

Элонгин (SIII) впервые был выделен из ядер печени крыс в виде белкового комплекса, состоящего из трех субъединиц А, В и С с молекулярными массами ~150, 18 и 15 кДа соответственно. Транскрипционная активность элонгина (SIII) целиком ассоциирована с его А-субъединицей, а две другие служат регуляторными и после образования стабильного димера оказывают сильное стимулирующее действие на транскрипционную активность А-субъединицы. Собственно стимулятором активности элонгина А является элонгин С, тогда как элонгин В, гомологичный убиквитину, не взаимодействует стабильно с элонгином А в отсутствие элонгина С и выполняет шапероноподобную функцию при сборке всего комплекса элонгина (SIII). На особую роль элонгина (SIII) в регуляции экспрессии генов указывает тот факт, что у человека он известен как потенциальная мишень действия продукта антионкогена (гена-супрессора опухолей) von Hippel-Lindau (VHL), мутации в котором ассоциированы с возникновением многих видов рака у человека. Белок VHL специфически взаимодействует с комплексом элонгина ВС, препятствуя

его связыванию с элонгином А. При этом мутации в антионкогене, сопровождающие онкологические заболевания, уменьшают прочность взаимодействия мутантного белка с элонгином ВС.

Ген фактора элонгации *ELL* (eleven–nineteen lysine-rich leukemia) человека, локализованный на хромосоме 19 (19p13.1), первоначально был обнаружен в связи с его частыми транслокациями в ген *MLL* (mixed lineage leukemia) на хромосому 11 (11q23) при острых миелоидных лейкозах. Предполагают, что продукт гена *MLL* участвует в регуляции транскрипции гомеостатических генов. В результате транслокации образуется "онкоген", кодирующий гибридный белок, который образован почти полным полипептидом *ELL*, объединенным с N-концевой частью белка *MLL*. Роль белка *ELL* в развитии лейкозов неясна, поскольку в настоящее время обнаружены шесть других генов, претерпевающих транслокацию в то же самое место на хромосоме 11, которые ассоциированы с лейкозами с различными клиническими проявлениями, характер которых зависит от природы транслоцируемого гена.

Терминация транскрипции. Прекращение синтеза РНК под действием РНК-полимеразы и освобождение РНК из транскрипционного комплекса происходят в конце транскрипционных единиц на особых участках ДНК - терминаторах транскрипции. Терминаторы транскрипции, функционирующие с разной эффективностью, могут находиться и внутри транскриптов. Такие терминаторы являются мощными факторами, регулирующими уровень транскрипции (и других этапов экспрессии) соответствующих генов. Для осуществления терминации транскрипции на некоторых терминаторах РНК-полимераза не требует дополнительных белковых факторов, тогда как другие терминаторы в их отсутствие не функционируют.

Терминация транскрипции у бактерий. Типичные терминаторы, не требующие для своего распознавания РНК-полимеразой *E. coli* дополнительных белковых факторов, содержат GC-богатый участок, обладающий центральной симметрией, вслед за которым располагается последовательность нуклеотидов, состоящая из выстроенных подряд четырех–восьми остатков А, в матричной цепи ДНК. Транскрипция завершается на конце этой олиго(А)-последовательности или же на следующем за ней нуклеотиде. Предполагается, что после прохождения РНК-полимеразой GC-богатого

участка ДНК с центральной симметрией в этом месте РНК образуется шпилька, что приводит к разрушению ДНК–РНК-гибрида в транскрибирующем комплексе.

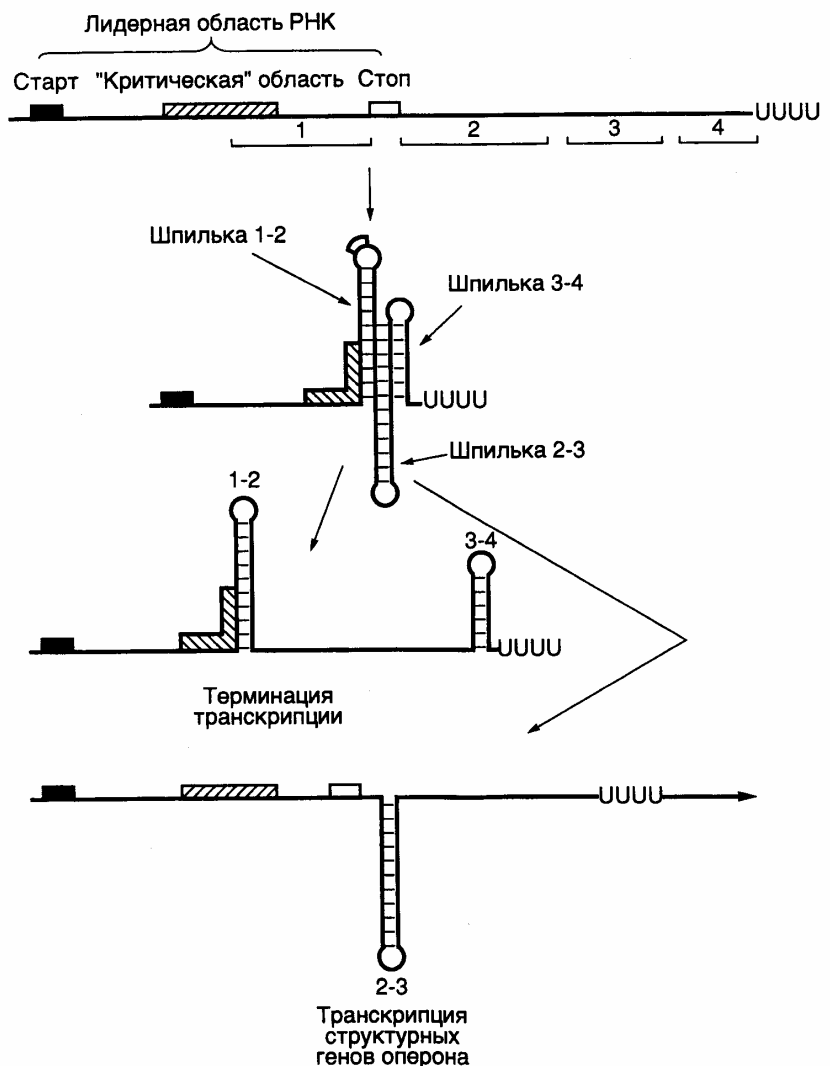


Рис. 1.9. Атенуатор триптофанового оперона *E. coli* и его функционирование

Изображены альтернативные вторичные структуры мРНК в районе аттенуатора, образование которых сопровождается прекращением транскрипции или распространением в область структурных генов оперона

Оставшаяся часть ДНК–РНК-гибрида нестабильна и легко плавится, поскольку образована 3'-концевой олиго(U)-последовательностью РНК и олиго(dA)-последовательностью терминатора. Первым из комплекса освобождается РНК-продукт, а затем минимальный фермент РНК-полимеразы.

После объединения со свободным σ -фактором образовавшийся холофермент РНК-полимеразы вступает в новый цикл транскрипции. Эффективность терминции транскрипции на таком терминаторе зависит от стабильности терминаторной шпильки РНК: мутации, нарушающие комплементарное спаривание оснований в шпильке, ослабляют терминцию, а мутации, восстанавливающие комплементарность, ее усиливают.

Кроме вышеописанного, у *E. coli* обнаружены терминаторы транскрипции, распознаваемые РНК-полимеразой только в присутствии белкового фактора терминции ρ , механизм действия которого хорошо изучен. Этот белок с молекулярной массой 46 кДа обладает РНК-зависимой нуклеозидтрифосфатазной активностью, которая необходима для его функционирования при терминции. Кроме того, для него характерна РНК:ДНК-хеликазная активность. Установлено, что фактор ρ связывается с растущей цепью РНК в особых неструктурированных участках, называемых *rut-сайтами* (rut sites, от англ. rut – колея, выбоина), до того, как РНК-полимераза достигает терминатора. В местах ρ -зависимой терминции транскрипции РНК-полимераза прекращает элонгацию. Считается, что роль ρ -фактора заключается в вытеснении РНК из транскрипционного комплекса во время таких пауз.

E. coli и другие бактерии имеют еще один тип регулируемых терминаторов транскрипции, называемых *аттенюаторами*. Впервые обнаружен и лучше других изучен аттенюатор триптофанового оперона, контролирующего биосинтез Тгр в бактериальных клетках (см. рис. I.9). В условиях избытка внутриклеточного Тгр девять из десяти молекул РНК-полимеразы, начавших транскрипцию триптофанового оперона, прекращают синтез РНК на аттенюаторе, расположенном на расстоянии в 180 п.о. от точки инициации транскрипции. В результате в основном происходит синтез коротких РНК той же длины, называемых лидерными. При уменьшении содержания Тгр в клетках доля молекул РНК-полимеразы, преодолевающих аттенюатор, возрастает, что, в конечном счете, сопровождается увеличением внутриклеточного содержания ферментов биосинтеза Тгр. Перед аттенюатором находятся несколько участков ДНК, последовательности которых обладают центральной симметрией. Это приводит к тому, что лидерная РНК, включающая в себя последовательности, комплементарные таким участкам, способна

образовывать структуры типа шпильки в разных сочетаниях, которые исключают друг друга. Например, если получена шпилька 2/3, то шпильки 1/2 и 3/4 сформироваться уже не могут. К аналогичным результатам приводит и обратное развитие событий. Шпилька 3/4 является терминаторной, присутствующей в ρ -независимых терминаторах. За ней в лидерной РНК располагается последовательность олиго(U). Поэтому образование шпильки 3/4 сопровождается терминацией транскрипции на аттенуаторе и освобождением лидерной РНК из транскрипционного комплекса. Формирование альтернативных шпилек зависит от положения рибосом, транслирующих лидерную РНК с образованием лидерного пептида, в котором присутствуют два остатка Trp подряд. В условиях недостатка Trp рибосома в процессе синтеза лидерного пептида останавливается на соответствующих кодонах лидерной РНК, прикрывая собой последовательность 1, что препятствует формированию шпильки 1/2, так как образуется шпилька 2/3. В соответствии с этим терминаторная шпилька не может сформироваться, транскрипция не прерывается на аттенуаторе и РНК-полимераза переходит в область структурных генов оперона. Если недостаток триптофана не приводит к прекращению трансляции лидерной РНК, рибосома проходит критический участок лидерной РНК, препятствуя формированию шпильки 2/3, и образуется терминаторная шпилька 3/4, что сопровождается терминацией транскрипции на аттенуаторе.

Терминация транскрипции у эукариот. У эукариот обнаружены три фактора терминации транскрипции, необходимых для освобождения РНК-полимераз из транскрипционных комплексов на терминаторах – по одному для РНК-полимераз I, II и III. Белок N-TEF дрожифилы индуцирует освобождение транскриптов, синтезированных РНК-полимеразой II, и при его функционировании происходит расщепление АТФ. У дрожжей белковый фактор Reb-1 связывается с природными терминаторами транскрипции на ДНК, обеспечивая как остановку элонгирующей РНК-полимеразы I на этих терминаторах, так и последующее освобождение РНК из транскрипционных комплексов. Удаление в результате делеции из рибосомной транскрипционной единицы Reb-1-связывающего сайта нарушает правильное образование 3'-концов рРНК *in vivo*. Мышиный фактор TTF-1, который также является ДНК-связывающим белком, необходим для правильной терминации транскрипции

РНК-полимеразой I в клетках этих животных. У них же обнаружен LA-белок, специфически взаимодействующий с РНК, функционирование которого требуется для образования транскриптов полной длины под действием РНК-полимеразы III, что происходит в результате освобождения РНК из транскрипционных комплексов и реинициации транскрипции.

Рассмотрим подробнее механизм терминации транскрипции Pol I. У мышей терминация транскрипции рДНК происходит на 565 п.о. ниже кодирующей части гена 28S РНК. 3'-Конец терминирующего транскрипта картирован за 21 п.о. перед 18-нуклеотидным повтором, названным Sal-боксом (AGGTCGACCAGA/TT/ANTCCG), который входит в состав терминатора транскрипции. Десять таких повторов, фланкированных протяженными кластерами пиримидиновых оснований, локализованы в нетранскрибируемых спейсерах рДНК. У человека длина повтора составляет 11 п.о. (GGGTCGACCAG), и его последовательность соответствует таковой 5'-концевой части мышинного повтора. Последовательности, фланкирующие повторы, оказывают влияние на точность и эффективность терминации транскрипции, а функционирование всего терминатора зависит от его ориентации на ДНК.

С помощью мутационного анализа и футпринтинга было установлено, что фактор терминации транскрипции TTF-I взаимодействует с Sal-боксом и останавливает элонгирующую Pol I. Хотя все факторы терминации транскрипции, обсуждавшиеся выше, распознают разные последовательности нуклеотидов, для них характерно наличие в С-концевых частях двух ДНК-связывающих доменов длиной в 80 аминокислот каждый, гомологичных ДНК-связывающей последовательности онкобелка c-Myc. Хотя более половины полипептидной цепи с N-конца TTF-I могут быть удалены без потери его функций, одних лишь ДНК-связывающих доменов недостаточно для обеспечения белком терминации транскрипции, и для этого требуются прилегающие последовательности аминокислот. Факторы терминации транскрипции TTF-I мышей и Rib-1 дрожжей могут прекращать элонгацию цепей РНК на любой из этих ДНК. Это указывает на высокую эволюционную консервативность механизма терминации транскрипции Pol I.

Связываясь с последовательностями терминаторов, TTF-I изгибает молекулу ДНК и вызывает задержку элонгирующего транскрипционного

комплекса на терминаторе. Предполагается, что в этот момент происходит конформационное изменение молекулы Pol I, что ослабляет взаимодействие компонентов комплекса друг с другом (как это имеет место в случае РНК-полимеразы *E. coli* на ρ -независимых терминаторах транскрипции). Окончательный распад комплекса и освобождение Pol I, а также синтезированной молекулы РНК происходит лишь в присутствии дополнительного фактора PTRF (polymerase and transcript release factor), который контактирует как с Pol I, так и с TTF-I.

Функциональная роль фактора TTF-I не ограничивается лишь участием в терминации транскрипции. Один из терминаторов транскрипции рДНК, так называемый T_0 , расположен за 170 п.о. перед точкой инициации транскрипции генов рРНК. Взаимодействующий с T_0 фактор TTF-I сильно стимулирует транскрипцию генов рРНК, вызывая перестройку структуры хроматина в окрестностях соответствующего промотора. Об изменениях структуры хроматина во время транскрипции см. следующий раздел 2.1.4.

Хотя для каждой из форм РНК-полимераз обнаружен свой специфический белковый фактор, необходимый для правильного освобождения транскриптов из элонгирующих комплексов, этим не ограничиваются механизмы, обеспечивающие терминацию транскрипции у эукариот. Действительно, одним из основных факторов терминации транскрипции у этой группы организмов является сложный белковый комплекс, обеспечивающий процессинг 3'-концевых последовательностей у предшественников мРНК, синтезируемых РНК-полимеразой II (см. ниже). В этом случае терминация транскрипции тесно сопряжена с процессингом пре-мРНК пока неизвестным молекулярным механизмом.

Терминация транскрипции митохондриальных ДНК человека, так же как и терминация синтеза рРНК, требует участия специального белкового фактора. В этом случае фактор MTTERM может ускорять терминацию транскрипции *in vivo* митохондриальной РНК-полимеразой, а также ферментами бактериального и фагового происхождения. В отличие от терминаторов генов рРНК, митохондриальные сигналы терминации транскрипции в комплексе с фактором MTTERM распознаются молекулами РНК-полимераз в обеих ориентациях, что делает возможной терминацию транскрипции на H- и L-цепях мтДНК. Вследствие этого один общий терминатор может обеспечивать

сбалансированное образование продуктов транскрипции с обоих противоположно направленных митохондриальных промоторов.

2.1.4. Хроматин во время транскрипции

В эукариотических клетках матрицей для РНК-полимераз служит ДНК, находящаяся в составе хроматина. Из общих соображений белки нуклеосом и более высокоорганизованного хроматина должны быть препятствием для образования инициационного комплекса и перемещения транскрипционного комплекса вдоль такой матрицы. Однако *in vivo* эти препятствия в соответствующих условиях легко преодолеваются. Исследование транскрипции хроматина в настоящее время еще далеко от своего завершения, а основные результаты получены в опытах *in vitro*. Тем не менее, в последнее время, благодаря взаимно дополняющим друг друга биохимическим и генетическим данным, становятся более ясными основные механизмы, обеспечивающие этот повсеместно распространенный процесс. Результаты такого рода исследований показывают, что наличие нуклеосом предотвращает неспецифическую транскрипцию инактивированных генов и является одним из необходимых условий их правильной экспрессии. Рассмотрим ряд современных моделей, в которых делается попытка объяснения таких механизмов.

Нуклеосомы и инициация транскрипции. Результаты многочисленных биохимических и генетических экспериментов показывают, что присутствие нуклеосом в промоторных участках генов, как правило, ингибирует транскрипцию. В частности,, установлено, что пространственное расположение последовательностей нуклеотидов в двух витках ДНК, намотанной вокруг гистонового октамера нуклеосомы, несовместимо со сборкой стабильного инициационного комплекса. Следовательно, для образования функционально активного инициационного комплекса, в состав которого входят РНК-полимераза и факторы транскрипции, необходимо локальное разрушение нуклеосомной структуры хроматина в окрестностях промотора и регуляторных элементов. При этом реализуются две стратегии: непрерывное существование участка ДНК в виде свободной от нуклеосом последовательности нуклеотидов и индуцированное разрушение нуклеосом. Первый механизм функционирует на промоторах конститутивно транскрибируемых эукариотических генов и

обеспечивается белковыми факторами, которые разрушают имеющуюся нуклеосомную структуру данного участка ДНК или препятствуют ее образованию. Описаны, по крайней мере, три способа осуществления этого механизма: 1) факторы транскрипции успевают взаимодействовать с реплицируемой ДНК до сборки нуклеосом; 2) факторы связываются с соответствующими участками ДНК, содержащими нуклеосомы, и их дестабилизируют; 3) специализированные белки разрушают нуклеосомную структуру в области промоторов неэкспрессирующихся генов. Все эти механизмы могут функционировать как по отдельности, так и в различных сочетаниях.

Вновь синтезируемая эукариотическая ДНК обладает повышенной чувствительностью к нуклеазам, что указывает на ее более "открытую" структуру по сравнению со структурой в сформированном хроматине интерфазных ядер. Это может отражать наличие промежуточных стадий в сборке нуклеосом или формировании структур хроматина более высокого порядка. Считается, что сборка нуклеосом в реплицирующейся ДНК происходит в два этапа. Вначале гистоны H3 и H4 доставляются к ДНК с помощью фактора сборки хроматина CAF-I (chromatin assembly factor I), состоящего из трех субъединиц с молекулярными массами 150, 60 и 50 кДа. Вновь синтезируемые гистоны H3 и H4 связываются первыми двумя субъединицами, из которых полипептид с молекулярной массой 150 кДа обладает сильно заряженным доменом, а другой содержит в своем составе WD-повтор (где W и D – аминокислоты Trp и Asn в соответствии с однобуквенным обозначением). На втором этапе к строящимся нуклеосомам добавляются гистоны H2A и H2B, что завершает формирование ядерных частиц нуклеосом. Исследования *in vitro* показали, что тетрамеры H3/H4 не исключают взаимодействия факторов транскрипции с соответствующими участками ДНК, как это делают зрелые нуклеосомы. С другой стороны, *нуклеоплазмин*, связывающий димеры H2A/H2B, стимулирует взаимодействие различных факторов с нуклеосомами (например GAL4, USF или Sp1). Кроме того, незрелый хроматин характеризуется пониженным содержанием линкерного гистона H1, присутствие которого стабилизирует нуклеосомы и структуры хроматина высшего порядка.

Конкурентные отношения между активацией транскрипции и созревaniem хроматина во время клеточного цикла были продемонстрированы *in vivo* для

дрожжевых генов, локализованных вблизи теломерных участков хромосом. У таких генов может иметь место мозаичный эффект положения (см. раздел 3.2.4). Например, транслокация гена *Ura3* в область теломеры часто подавляет его транскрипцию, которая может быть возобновлена под действием белка-трансактиватора транскрипции Prg1, но только в фазе G2/M клеточного цикла. Именно в это время происходит полное созревание вновь образуемого хроматина у большинства эукариотических организмов. Однажды установившись, активированное или репрессированное состояние гена вблизи теломерных участков хромосом поддерживается на протяжении многих клеточных делений. Эти эксперименты показывают, что во время сборки хроматина имеется возможность перепрограммирования компетентности генов в отношении транскрипции и регуляторная структура хроматина является наследуемой в ряду клеточных поколений.

Промоторы, активируемые через разрушение нуклеосомной структуры непосредственно факторами транскрипции, характерны для генов теплового шока дрозофилы. В поддержании открытой структуры ДНК в этом случае участвуют основные факторы транскрипции, а также GAGA-фактор, которые взаимодействуют с промотором в окрестностях ТАТА-бокса и точки инициации транскрипции. Такое взаимодействие обеспечивает открытое состояние вышерасположенного регуляторного элемента теплового шока. При индуцированном механизме разрушения нуклеосомной структуры ДНК в окрестностях промотора перед активацией гена нуклеосомы присутствуют как в вышерасположенных регуляторных последовательностях ДНК, так и в самом промоторе. Во время индукции транскрипции такого гена регуляторные факторы, связываясь с ДНК, прямо или косвенно вызывают разрушение нуклеосомной структуры соответствующих участков ДНК. Аналогичная стратегия активации промоторов реализуется также и в генах, регулируемых глюкокортикоидами. Для активации промоторов, структурированных в нуклеосомы, требуется несколько этапов. Вначале регуляторные факторы, структура которых более подробно обсуждается в разделе 3.2.2, взаимодействуют своими ДНК-связывающими доменами с соответствующей регуляторной последовательностью, расположенной выше промотора, что сопровождается вытеснением части или всех гистонов нуклеосом этой последовательности. Активирующие домены белковых факторов транскрипции

далее индуцируют освобождение гистонов с основного промотора, что сопровождается образованием инициационного комплекса с участием РНК-полимеразы и основных факторов транскрипции. Сборка транскрипционного комплекса приводит к вытеснению еще одной порции гистонов с промотора.

Модификация структуры нуклеосом через ацетилирование гистонов. Рассмотренный пример показывает, что не только ДНК-связывающие домены факторов транскрипции адаптируют структуру хроматина для нужд синтеза РНК, но и их активирующие домены участвуют в разрушении структуры нуклеосом, расположенных на ДНК вблизи мест связывания факторов. Активирующие домены таких факторов взаимодействуют не только с основными факторами транскрипции, такими как TFIID, но и с коактиваторами транскрипции, которые не входят в состав основного транскрипционного комплекса, хотя и обеспечивают его функциональную связь с удаленными регуляторными последовательностями энхансеров и сайленсеров (подробнее см. раздел 3.2.2). До недавнего времени оставалось совершенно непонятным, каким образом активирующие домены факторов транскрипции могут оказывать влияние на структуру близлежащего хроматина. Этот вопрос начал проясняться после недавнего клонирования генов *ядерных и цитоплазматических ацетилаз гистонов* (НАТ – histone acetyltransferase) и изучения свойств этих ферментов в очищенном состоянии.

Ацетилирование гистонов происходит на посттрансляционном уровне по специфическим остаткам Lys в N-концевых частях их полипептидных цепей, которые расположены на поверхности нуклеосомных частиц (см. табл. I.2). Эта посттрансляционная модификация уменьшает суммарный положительный заряд ядерных частиц нуклеосом и ослабляет взаимодействие плеч гистонов с ДНК. Хотя при этом сохраняется общая структурная целостность нуклеосом, их конформация может изменяться, что ингибирует *in vitro* зависимое от ионной силы образование структур хроматина более высокого порядка, связанное с подавлением транскрипции. Кроме того, ацетилирование гистонов нарушает специфические взаимодействия между их плечами и некоторыми регуляторными белками-репрессорами. В соответствии с этим, уже давно была обнаружена корреляция между повышенным уровнем ацетилирования гистонов и усилением транскрипции определенных участков генома, а пониженный уровень ацетилирования гистонов был связан с молчащими генами и

гетерохроматином. Однако оставался непонятным механизм избирательного ацетилирования гистонов на промоторах активируемых генов.

Ацетилазы гистонов эукариот разделяют на две основные группы – НАТ А и НАТ В. Для ацетилаз НАТ А характерна ядерная локализация. Эта группа ферментов участвует в посттрансляционном ацетилировании гистонов ядерных частиц нуклеосом и оказывает прямое влияние на транскрипцию. Цитоплазматические ацетилазы НАТ В преимущественно модифицируют молекулы гистонов, находящиеся в свободном состоянии, и участвуют в их доставке к реплицируемой ДНК. Для этих двух систем ацетилирования характерна разная субстратная специфичность. Если ацетилазы НАТ А ацетилируют все четыре гистона ядерных частиц нуклеосом, то ацетилазы НАТ В модифицируют, прежде всего, гистоны H3 и H4 по другим остаткам Lys.

Дрожжевая цитоплазматическая ацетилаза гистонов типа В, кодируемая геном *HAT1*, по-видимому, обеспечивает направленную доставку вновь синтезированных гистонов в ядра, где характер их ацетилирования может быть изменен ядерной ацетилазой НАТ А. Последовательность нуклеотидов клонированного гена ацетилазы НАТ А *Tetrahymena* неожиданно оказалась высокоомологичной последовательности гена известного коактиватора транскрипции дрожжей *Gcn5*, продукт которого (Gcn5), как выяснилось, также обладает активностью ацетилазы гистонов. Эта находка явилась первым указанием на то, что ацетилирование гистонов может быть одной из причин, а не следствием активации генов. Белок Gcn5 образует тримерный комплекс с двумя другими белками – Ada2 и Ada3, которые необходимы для активации генов кислыми факторами транскрипции Gal4–Vp16 и Gen4. При этом комплекс Gcn5–Ada2–Ada3 образует прямые контакты и с *транс*-действующими, и с основными факторами транскрипции в составе транскрипционного комплекса. Это позволяет объяснить механизм направленной доставки ацетилазы гистонов к активируемому промотору соответствующими специфическими белок-белковыми взаимодействиями. Интересной находкой этой серии экспериментов явилось понимание двойной роли комплекса Gcn5–Ada в активации транскрипции. Во-первых, он ацетилирует гистоны нуклеосом промотора, что повышает его доступность для основных факторов транскрипции и других регуляторных белков. Во-вторых, он непосредственно контактирует с этими факторами, облегчая образование прединициационного

комплекса и стабилизируя его.

Таким образом, индивидуальные факторы транскрипции могут выполнять несколько разных функций в регуляции синтеза РНК, определяемых их выраженной доменной структурой. Как уже упоминалось, для полипептидных цепей большинства факторов транскрипции характерно наличие, по крайней мере, ДНК-связывающего и активирующего доменов. Из вышеизложенного следует, что ДНК-связывающие домены этих белков могут непосредственно изменять структуру нуклеосом в окрестностях промоторов, а активирующие домены не только контактируют с основными факторами транскрипции и стабилизируют их связь с промоторами, но и ассоциированы с активностями, модифицирующими структуру хроматина, что необходимо для эффективной инициации транскрипции и освобождения промотора элонгирующим транскрипционным комплексом. И все эти активности у дрожжей присущи одному сравнительно небольшому белковому комплексу Gcn5–Ada.

Деацетилазы гистонов. Стационарный уровень ацетилирования гистонов хроматина поддерживается в результате координированного действия НАТ и деацетилаз гистонов (histone deacetylase – Hd), наиболее изученными из которых являются деацетилазы гистонов дрожжей, дрозофилы и человека. Быстрая очистка деацетилаз достигается с помощью аффинной хроматографии, в которой используется в качестве лиганда иммобилизованный высокоаффинный ингибитор *трапоксин*. При этом деацетилазы гистонов дрожжей выделяются в виде двух высокомолекулярных комплексов - HdA (350 кДа) и HdB (600 кДа). HdA может деацетилировать все четыре гистона и сильно ингибируется *трихостатином А*, тогда как HdB в 10 раз менее чувствительна к действию этого ингибитора. Очищенная HdA содержит в своем составе четыре полипептида, два из которых с молекулярными массами 75 и 71 кДа кодируются генами *HDA1* и *HDA3* соответственно. Деацетилаза HdB содержит в своем составе белок Rpd3, функционирование которого, как это было показано генетическими методами, необходимо для достижения не только полного подавления, но и полной активации экспрессии большого числа генов. В состав деацетилазы HdA входит Rpd3-подобный белок; такие белки представлены у дрожжей целым семейством, насчитывающим, по крайней мере, четыре члена. Инактивация деацетилаз HdA и HdB с помощью делеций в соответствующих генах приводит к гиперацетилированию гистонов H3 и H4 *in vivo*.

Как уже упоминалось, гиперацетилирование гистонов, как правило, коррелирует с активацией транскрипции определенных генетических локусов. Однако повышенный уровень ацетилирования гистонов может сопровождаться подавлением экспрессии генов, локализованных в теломерных участках хромосом дрожжей. Аналогичную ситуацию наблюдали у дрозофилы: инактивация гомологичного гена деацетилазы приводила к усилению эффекта положения гена *white*, транслоцированного в область центромерного гетерохроматина. Эти странные эффекты могут быть связаны с нарушением специфичности ацетилирования гистонов в условиях их гиперацетилирования при инактивации генов деацетилаз. Действительно, для подавления транскрипции у дрожжей необходимо ацетилирование единственного аминокислотного остатка (Lys-12 в гистоне H4). Кроме того, те же гистоны, ацетилированные по тому же самому положению, были обнаружены в транскрипционно неактивном β -гетерохроматине поличенных хромосом дрозофилы.

Как и в случае ацетилаз гистонов, специфический характер действия деацетилаз обеспечивается путем образования комплексов с белками-репрессорами и корепрессорами, которые распознают конкретные последовательности ДНК и друг друга. Ядерные рецепторы тиреоидных гормонов особенно хорошо иллюстрируют функционирование такого механизма. В отсутствие лиганда они взаимодействуют с комплексом репрессор–деацетилаза, что приводит к подавлению транскрипции соответствующих генов, тогда как под действием гормона рецепторы приобретают способность образовывать комплекс с коактиватором транскрипции и ацетилазой и стимулировать синтез РНК.

Другие специализированные белки, изменяющие структуру хроматина. В дополнение к вышеупомянутым белкам в настоящее время выделены и охарактеризованы два других специфических белковых комплекса, обеспечивающих изменение структуры нуклеосом во время транскрипции. Комплекс Swi/Snf был впервые обнаружен в клетках дрожжей генетическими методами как позитивный регулятор транскрипции большого числа генов, экспрессия которых контролируется разными механизмами. Похожий по механизму действия комплекс NURF (nucleosome remodeling

Таблица I.6

**Субъединичный состав и свойства белковых комплексов
Swi/Snf и NURF**

Фактор	Субъединицы	Общая молекулярная масса, кДа	Активность
Swi/Snf	Swi1	2000	ДНК- зависимая АТРаза
	Swi2 (Snf2)		
	Swi3		
	Snf3		
	Snf5		
	Snf6		
	p78		
	p68		
	p50		
	p47		
	p25		
NURF	215 кДа	500	АТРаза, зависимая от нуклеосом
	140 кДа		
	55 кДа		
	38 кДа		

factor) был выделен из эмбрионов дрозофилы (табл. I.6).

Оба комплекса состоят из нескольких субъединиц, обладают АТРазной активностью и оказывают влияние на контакты между гистонами и ДНК. Однако для этих комплексов характерно отсутствие общих компонентов, и они различаются по механизмам стимуляции АТРаза. Такие различия между комплексами позволяют предполагать, что они действуют на разные субстраты и могут функционировать независимо друг от друга.

Первые указания на то, что факторы Swi/Snf участвуют в изменении структуры хроматина, были получены генетическими методами. Индивидуальные мутации *SWI* или *SNF*, ингибирующие экспрессию ряда генов, супрессировались мутациями, повышающими внутриклеточный уровень гистонов. Комплексы Swi/Snf, выделенные из клеток дрожжей и человека, содержали, по крайней мере, десять различных белков. Первые шесть компонентов, перечисленные в табл. I.6, были идентифицированы как продукты конкретных генов, тогда как другие (p78–p25) – только как полипептиды указанной молекулярной массы. Комплекс Swi/Snf стимулирует *in vitro* АТРа-зависимое связывание нуклеосомами *транс*-действующих факторов транскрипции. Мутационные изменения консервативных участков гистонов H3 и H4, необходимых для формирования у них правильной пространственной структуры, делают транскрипцию независимой от фактора Swi/Snf *in vivo*.

Недавно комплекс Swi/Snf удалось выделить в составе холофермента РНК-полимеразы II, так что этот комплекс может оказаться неотъемлемой частью транскрибирующего фермента и присутствовать в инициационных комплексах всех промоторов. Однако в таком случае остается не совсем понятным, почему мутации в генах *SWI/SNF* оказывают влияние на экспрессию лишь некоторых генов. Поскольку белки Swi/Snf ассоциированы с медиаторным комплексом, который, в свою очередь, связан с СТД-доменом РНК-полимеразы II и обеспечивает ответ РНК-полимеразы на действие регуляторных факторов, последние могут по-разному реагировать на присутствие в комплексе мутантных белков Swi/Snf, что и может быть причиной дифференциального ответа конкретных промоторов на мутации.

Белковый комплекс NURF впервые был описан как кофактор уже упоминавшегося выше фактора транскрипции GAGA, который необходим для активации промотора гена теплового шока дрозофилы *hsp70*. *In vivo* в этом

промоторе была обнаружена последовательность длиной в 200–300 нуклеотидов с повышенной чувствительностью к ДНКазе, в состав которой входят TATA-бокс и сайты связывания факторов GAGA и HSF (heat shock factor – фактор теплового шока). Гиперчувствительность промотора к действию ДНКазы можно было моделировать в бесклеточных экстрактах как до, так и после сборки хроматина путем добавления GAGA-фактора и АТФ. Поскольку фактор GAGA сам по себе не обладает АТФазной активностью, в результате очистки белков с этой активностью и был идентифицирован комплекс NURF. Этот комплекс может разрушать нуклеосомы или изменять их положение в промоторе гена *HSP70* и в отсутствие фактора GAGA. Однако присутствие NURF стимулирует связывание этого фактора со своим сайтом на промоторе, что, в свою очередь, ускоряет перестройку нуклеосом в этом участке гена. АТФазная активность NURF не стимулируется свободной ДНК или гистонами, однако усиливается в присутствии интактных нуклеосом, что отличает этот фермент от АТФазы Swi2. При микросеквенировании пептида с молекулярной массой 140 кДа в нем был обнаружен АТФазный домен, гомологичный таковому Swi2, однако это оказалось единственной гомологией между двумя комплексами. Следует еще раз подчеркнуть, что отсутствие у них общих субъединиц и существенной гомологии указывает на возможность независимого функционирования этих комплексов и действия на разные субстраты.

Специфичность функционирования АТФ-зависимых белковых комплексов, изменяющих структуру нуклеосом, предполагает их точную доставку в нужные места хроматина. Как и в уже рассмотренном случае ацетилаз и деацетилаз гистонов, специфический характер взаимодействия комплексов с ДНК обеспечивается дополнительными белками. Выше упоминалось о том, что комплекс Swi/Snf входит в состав холофермента РНК-полимеразы II, что обеспечивает его доставку к соответствующим промоторным последовательностям. Кроме того, было показано, что этот комплекс может взаимодействовать с рецептором глюкокортикоидов и в таком виде оказывать влияние на структуру нуклеосом. Формирование таких комплексов, по-видимому, является одним из существенных моментов активации рецепторов глюкокортикоидов. Учитывая обсуждавшееся выше взаимодействие рецепторов стероидных гормонов с ацетилазами/деацетилазами гистонов, можно полагать,

что специфическое изменение структуры хроматина является общим механизмом, посредством которого рецепторы оказывают влияние на экспрессию соответствующих генов.

Нуклеосомы при элонгации синтезируемой РНК. Механизмы, обеспечивающие элонгацию транскрипции на нативном хроматине, не совсем понятны. Поскольку РНК-полимеразы прокариот, в частности фагов SP6 и T7, обладают способностью транскрибировать хроматин *in vitro*, создавалось впечатление, что для прохождения РНК-полимеразами нуклеосом хроматина во время элонгации РНК не требуются дополнительные факторы. Тем не менее, нуклеосомы ингибируют транскрипцию хроматина *in vitro* на стадии элонгации эукариотическими РНК-полимеразами II и III, что не наблюдается *in vivo*. Одной из гипотез, объясняющих процесс элонгации транскрипции на хроматине, является *модель двойных суперскрученных доменов*. В соответствии с этой моделью предполагается, что транскрибирующая РНК-полимераза индуцирует в ДНК образование локальных суперскрученных доменов. Положительные супервитки образуются впереди элонгирующей РНК-полимеразы, а отрицательные – позади фермента. Закручивание ДНК вокруг гистонового октамера в нуклеосоме приводит к незначительным изменениям в параметрах двойной спирали ДНК и к образованию одного отрицательного супервитка в молекуле ДНК. Его формирование должно приводить к компенсаторной положительной сверхспирализации участков ДНК, прилегающих к нуклеосоме. Образование нуклеосом осуществляется предпочтительно на отрицательно сверхспирализованной ДНК, а ее положительная сверхспирализация сопровождается ослаблением структуры нуклеосом или их разрушением в присутствии дополнительных белковых факторов. Эти факты и лежат в основе обсуждаемой модели.

Таким образом, в соответствии с этой моделью, элонгирующая РНК-полимераза индуцирует впереди себя локальную положительную сверхспирализацию молекулы ДНК, что облегчает разрушение нуклеосом, находящихся в этой зоне. Повторное образование нуклеосом происходит в зоне отрицательно сверхспирализованной ДНК позади транскрибирующей РНК-полимеразы.

Конвергентный характер исследований функциональной структуры хроматина и транскрипции лишь иллюстрирует общую тенденцию развития

современной молекулярной биологии и генетики. Чем глубже становится понимание механизмов функционирования отдельных генетических систем клетки, тем яснее видится их взаимозависимость и полифункциональность. Высокоупорядоченные перестройки нуклеосом и хроматина сопровождают не только транскрипцию, но и репликацию, рекомбинацию и репарацию повреждений ДНК. В связи с этим проблема структуры хроматина и динамики ее изменений в клеточном цикле является одной из центральных в современной молекулярной генетике.

Концепция транскриптосомы. Как было показано выше, транскрипционный комплекс, в состав которого входит эукариотическая РНК-полимераза II, устроен весьма сложно. Появляется все больше данных в пользу того, что транскрипционный комплекс взаимодействует с другими крупными белковыми комплексами, участвующими, в частности, в разрушении нуклеосом и репарации ДНК. Полный размер образующегося при этом стабильного транскрипционного комплекса, содержащего более 70 полипептидов, приближается к размеру рибосомы. Такой колоссальный размер транскрипционного комплекса эукариот, вероятно, должен замедлять поиск путем линейной диффузии регуляторных последовательностей нуклеотидов транскрибируемой ДНК, на которых происходит инициация транскрипции. Обсуждается возможность того, что инициация транскрипции у эукариот осуществляется в специализированных надмолекулярных комплексах, специфически ассоциированных с ядерным матриксом, которые получили название *транскриптосом*. По крайней мере, один из белковых компонентов, входящих в состав холофермента POL II животных, YY1, оказался идентичным фактору NMP-1, ассоциированному с ядерным матриксом. Возможно, именно с участием этого белка происходит прикрепление транскриптосом к ядерным мембранам.

Подводя итог рассмотрению основных этапов транскрипции, необходимо отметить, что инициация синтеза РНК, элонгация транскриптов и терминация транскрипции являются очень сложно организованными процессами. Структуры матричной ДНК и растущих цепей РНК оказывают влияние на процесс освобождения промотора РНК-полимеразами, а также на свойства самих элонгирующих и терминирующих ферментов. При этом на инициацию, задержку и прекращение транскрипции, расщепление РНК и ее реитеративный синтез, а

также на саму терминацию оказывают действие многочисленные белковые факторы. Все это находит свое выражение в сложности регуляторных процессов, обеспечивающих координированную экспрессию генов на уровне транскрипции. Основные биохимические механизмы, контролирующие эти процессы, будут рассмотрены в соответствующих разделах книги.

2.2. Котранскрипционные и посттранскрипционные модификации РНК

Транскрипция у любого организма является первым этапом реализации генетической информации – экспрессии генов. Однако первичные транскрипты, как правило, представляют собой лишь предшественники зрелых мРНК – пре-мРНК, которые перед выполнением своих функций должны претерпеть многочисленные изменения, называемые *посттранскрипционными модификациями*. Кроме того, у эукариот зрелые мРНК должны быть доставлены от места их биосинтеза (ядра) к месту трансляции (цитоплазматическим рибосомам), т.е. экспортироваться из ядра в цитоплазму.

Одной из наиболее удивительных посттранскрипционных модификаций пре-мРНК является редактирование их первичной структуры. В результате посттранскрипционно изменяется смысл генетической информации, заключенной в соответствующем гене.

Посттранскрипционные модификации РНК особенно характерны для эукариотических организмов, у которых в силу мозаичной интрон-экзонной структуры их генов первичные транскрипты представлены гигантскими предшественниками, включающими в себя последовательности как экзонов, так и интронов. 5'-Конец предшественника мРНК чаще всего подвергается котранскрипционным модификациям, в результате которых к его 5'-концевому нуклеотиду особым образом присоединяется остаток гуанозина с образованием "шапочки" – *кэпа*. Эта котранскрипционная модификация создает условия для прохождения следующего этапа процессинга мРНК – *сплайсинга*, сопровождающегося вырезанием последовательностей интронов и объединением экзонов с образованием непрерывной кодирующей последовательности мРНК. Одновременно от 3'-конца путем эндонуклеазного расщепления отделяется избыточный фрагмент РНК, и к оставшейся части присоединяется поли(А)-последовательность. Эта совокупность реакций получила название *полиаденилирования* мРНК. После таких

котранскрипционных и посттранскрипционных модификаций пре-мРНК образовавшаяся зрелая, стабилизированная мРНК переносится из ядра в цитоплазму, часто к специфическому месту своей внутриклеточной локализации, где может быть депонирована или эффективно транслироваться рибосомами. Каждый из этапов посттранскрипционных модификаций может использоваться для регуляции уровня экспрессии соответствующих генов.

2.2.1. Процессинг РНК у бактерий

мРНК прокариот обычно являются полицистронными, т.е. включают в себя последовательности нуклеотидов нескольких генов одного оперона (рис. 1.10,а). Полицистронные мРНК бактерий при выполнении своих функций матричных РНК в трансляции не требуют разбиения на последовательности отдельных генов и могут транслироваться непосредственно рибосомами с образованием функционально активных белков. Исключением из правила являются полицистронные ранние мРНК нечетных Т-бактериофагов Т3 и Т7, которые после транскрипции *in vivo* расщепляются до моноцистронных под действием РНКазы III, специфически гидролизующей двухцепочечные РНК. Участие этого фермента в процессинге указывает на наличие характерной вторичной структуры РНК на границах транскриптов отдельных генов вышеупомянутых бактериофагов, что и было обнаружено после определения их первичной структуры.

РНКазы III участвуют также в процессинге предшественников рРНК у *E. coli*, поскольку 5S, 16S и 23S рРНК исходно синтезируются в составе общего первичного транскрипта. При этом в спейсерных участках 30S предшественников между последовательностями 16S и 23S рРНК расположены тРНК. Кроме того, у бактерий обнаружены транскрипты генов тРНК, содержащие до шести тРНК в составе одного предшественника. В процессинге предшественников тРНК у бактерий ключевую роль играет РНКазы Р. Для проявления нуклеазной активности у этого необычного фермента требуется присутствие небольшой РНК, прочно ассоциированной с его полипептидной цепью. Именно ей присуща собственная эндонуклеазная активность, что характерно и для других рибозимов (подробнее см. главу 9). Таким образом, зрелые молекулы рРНК и тРНК образуются в клетках как прокариот, так и эукариот в результате серии эндо- и экзонуклеазных

воздействий на их предшественники. Основной же посттранскрипционной модификацией является полиаденилирование их 3'-концевых последовательностей.

Полиаденилирование РНК у бактерий. Хотя поли(А)-полимераза *E. coli*, осуществляющая безматричный синтез поли(А) при наличии РНК-затравок, была очищена в 1962 г., поли(А)-РНК у бактерий не была обнаружена до 1975 г.

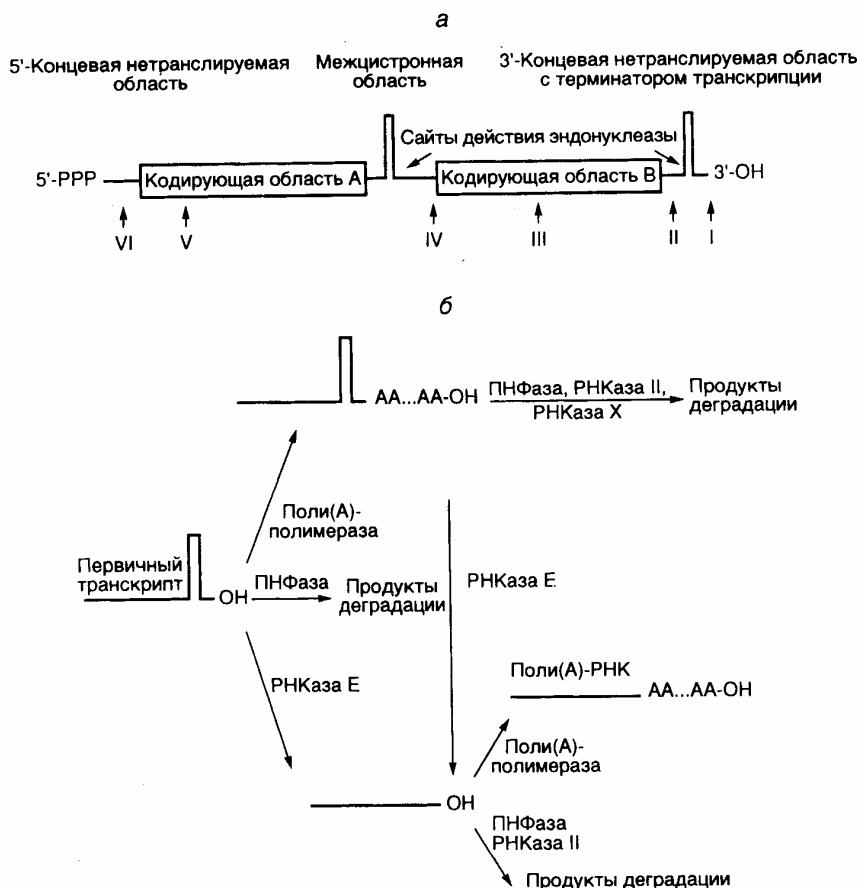


Рис. I.10. Локализация сайтов полиаденилирования и пути катаболизма бактериальных РНК

а – обобщенная структура бактериальной полицистронной мРНК и положение сайтов полиаденилирования в РНК классов I-VI; **б** – альтернативные пути катаболизма бактериальных мРНК с участием поли(А)-полимеразы, полинуклеотидфосфорилазы (ПНФазы) и различных РНКаз; РНККаза X – гипотетическая рибонуклеаза

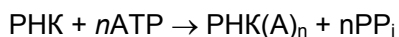
и долгое время после этого рассматривалась как исключение из общего

правила и биологический курьез. В настоящее время стало ясно, что полиаденилирование РНК у бактерий столь же обычно, как и у эукариот, и выполняет важные биологические функции. Поли(А)-последовательности бактериальных мРНК значительно короче соответствующих эукариотических. Их длина, в среднем, составляет всего 14–16 нуклеотидов (80–200 – у эукариот), а полиаденилированы лишь от 1 до 40% молекул мРНК каждого определенного вида в клетке (~100% в случае эукариот). Более подробное сравнение свойств поли(А)-последовательностей прокариот и эукариот проведено в разделе 2.2.3.

Классы поли(А)-содержащих РНК бактерий. В зависимости от расположения сайта присоединения поли(А)-последовательности все бактериальные РНК разделяют на шесть классов (см. рис. I.10,а). В РНК первого класса происходит полиаденилирование 3'-концевого нуклеотида, следующего сразу за терминаторной шпилькой ρ -независимого терминатора последней кодирующей области. У РНК второго класса терминаторная шпилька отщепляется посттранскрипционно РНКазой E, и к образующемуся 3'-концу присоединяется поли(А). Терминаторная шпилька отсутствует и у полиаденилированных мРНК третьего класса, образующихся с участием ρ -зависимых терминаторов транскрипции, благодаря особенностям структуры таких терминаторов. В укороченных мРНК четвертого класса отсутствует 3'-концевая кодирующая область, а поли(А) начинается непосредственно за межцистронной терминаторной шпилькой. В РНК пятого класса поли(А)-последовательности локализуются на концах укороченных цистронов, а у РНК шестого класса – на концах 5'-концевых некодирующих последовательностей. На основании структуры поли(А)⁺-мРНК делается вывод, что полиаденилирование у бактерий не зависит от присутствия в РНК специфических регуляторных сигналов, как это имеет место у эукариот, а определяется наличием у них свободных 3'-ОН-групп. Такие промежуточные мРНК разной длины, по-видимому, появляются в результате деградации или преждевременной терминации синтеза полноразмерных транскриптов.

Поли(А)-полимеразы. Полиаденилирование бактериальных мРНК катализирует поли(А)-полимераза (АТФ:полирибонуклеотид-аденилилтрансфераза), которая осуществляет независимое от матрицы

последовательное присоединение остатков аденилата к 3'-ОН-концам молекул РНК в соответствии со следующей реакцией:



У *E. coli* описаны две поли(А)-полимеразы. Поли(А)-полимераза I кодируется локусом *psnB*, первоначально идентифицированным в качестве контролирующего число копий плазмид в бактериальных клетках. Процессированная молекула представляет собой полипептид с молекулярной массой 52 кДа, не обладающий гомологией с соответствующими эукариотическими ферментами, однако содержащий сегмент, гомологичный части полипептидной цепи тРНК-нуклеотидилтрансферазы, которая осуществляет посттранскрипционное присоединение акцепторного ССА-тринуклеотида к молекулам тРНК. Даже умеренная сверхэкспрессия рекомбинантного гена *psnB* летальна для *E. coli*. Полная инактивация гена *psnB* с помощью делеций не сопровождается прекращением полиаденилирования мРНК, что указывает на присутствие в геноме *E. coli* второго аналогичного гена. Такой ген был обнаружен в виде открытой рамки считывания *f310*, кодирующей полипептид с молекулярной массой 36,3 кДа, обогащенный гидрофобными аминокислотами. Полипептид не обладает гомологией ни с одной из известных бактериальных, вирусных или эукариотических поли(А)-полимераз.

Возможная функциональная роль полиаденилирования РНК у бактерий. Несмотря на то что полиаденилирование РНК интенсивно исследуется более 25 лет, его функциональная роль даже у эукариот полностью не выяснена. Еще меньше известно о роли полиаденилирования РНК у бактерий. Однако из имеющихся данных становится ясно, что влияние полиаденилирования РНК на молекулярные процессы бактериальной клетки весьма разнообразно и распространяется, по крайней мере, на контроль числа копий бактериальных плазмид, стабильность мРНК и ее трансляцию.

Регуляция репликации плазмид через полиаденилирование антисмысловых РНК. Более подробно механизмы регуляции репликации бактериальных плазмид будут рассмотрены в разделе 4.2.2 на примере плазмиды ColE1. Здесь же лишь кратко отметим, что у некоторых групп плазмид, например ColE1 или pBR322, которая является ее производной, регуляция числа копий осуществляется с помощью антисмысловой РНК (РНК I),

образующей гибрид с РНК-праймером (РНК II), необходимым для инициации их репликации, и блокирует его функционирование. В соответствии с этим внутриклеточная концентрация РНК I является критическим параметром в регуляции репликации плазмид. Деградация РНК I инициируется отщеплением 5'-концевого пентануклеотида под действием РНКазы E (РНК I₅) и далее контролируется поли(А)-полимеразой – продуктом гена *psnB*. Полиаденилированная РНК I₅ обладает коротким временем полужизни, типичным для бактериальных мРНК (1–2 мин), тогда как немодифицированная РНК I₅ значительно более стабильна (время полужизни – >10 мин). Ускоренная деградация поли(А)-РНК I₅ инициируется полинуклеотидфосфорилазой. У поли(А)-РНК I₅ 3'-концевой нуклеотид находится в составе шпильки, что препятствует эффективному действию полинуклеотидфосфорилазы. Однако при наличии короткого поли(А)-хвоста, выступающего за пределы шпильки, РНК I₅ эффективно расщепляется ферментом по 3'-экзонуклеазному механизму. В разрушение поли(А)⁺-РНК I₅ вносят свой вклад и другие нуклеазы, включая РНКазу E и РНКазу III. Имеются данные и том, что полиаденилирование само по себе инактивирует РНК I₅, так как оно приводит к характерному изменению ее вторичной структуры.

Влияние полиаденилирования на время полужизни бактериальных мРНК. Описанное выше дестабилизирующее действие полиаденилирования на антисмысловые РНК распространяется и на некоторые бактериальные мРНК. Например, инактивация гена *psnB*, кодирующего поли(А)-полимеразу, с помощью делеций сопровождается значительным увеличением времени полужизни мРНК генов *lpp*, *trxA*, *ompA*, а также *rpsO* на фоне мутаций *pnp*, *rnb* и *rne*, инактивирующих гены полинуклеотидфосфорилазы и соответствующих рибонуклеаз. Поскольку эти результаты были получены в отсутствие РНКазы E, 3'-концы большинства бактериальных мРНК содержали шпильку, характерную для р-независимых терминаторов транскрипции (см. рис. I.10). Экзонуклеазное расщепление РНК под действием полинуклеотидфосфорилазы и РНКазы II затруднено при наличии такой шпильки и облегчается в присутствии выступающей поли(А)-последовательности (см. рис. I.10,б). Как показано на рисунке, первичный транскрипт в бактериальных клетках может или полиаденилироваться, или подвергнуться экзонуклеазному расщеплению, неэффективному без поли(А)-последовательности, при наличии которой в

процесс деградации активно включаются полинуклеотидфосфорилаза и РНКазы II. РНКазы E может отщеплять как полиаденилированную, так и неполиаденилированную шпильку. Такой незащищенный с 3'-конца транскрипт подвергается атаке 3'-экзонуклеаз, с которой конкурирует полиаденилирование, препятствующее деградации мРНК по этому механизму. Следовательно, полиаденилирование мРНК в бактериальных клетках может выполнять альтернативные функции: дестабилизировать транскрипты при наличии у них 3'-концевых шпилек и стабилизировать "линейные" формы мРНК. Предполагается участие в деградации поли(А)⁺-мРНК и неизвестной РНКазы X, поскольку этот процесс может происходить на фоне неактивных полинуклеотидфосфорилазы, РНКазы II и РНКазы E.

Изображенная на рис. 1.10 схема катаболизма первичных бактериальных транскриптов является упрощением, так как предполагает независимое действие нуклеаз. В последнее время, в соответствии с общей тенденцией, накапливаются данные о координированной работе большинства компонентов системы деградации бактериальных РНК в составе сложных мультиферментных комплексов, получивших название *деградосом*. Полное расщепление РНК с 3'-концевой шпилькой деградосомами требует расхода АТФ АТФ-зависимой РНК-хеликазой, а также участия белка DnaK, белков теплового шока, энолазы и гликолитического фермента с неизвестной функцией в метаболизме РНК.

Возможная роль полиаденилирования бактериальных мРНК в трансляции. Как будет видно из дальнейшего изложения (см. раздел 2.2.3), поли(А)-последовательности эукариотических РНК постоянно ассоциированы с жизненно важным поли(А)-связывающим белком РАВ, который участвует как в регуляции стабильности мРНК, так и их трансляции. Поиск белка с аналогичными функциями у бактерий привел к очистке рибосомного белка S1, который кооперативно связывается с поли(А)-последовательностями мРНК с константой ассоциации $3 \cdot 10^6 \text{ M}^{-1}$. Функциональная роль этого взаимодействия в настоящее время не ясна, однако предполагают, что оно может оказывать влияние на трансляцию. Возможно, белок S1 облегчает доставку мРНК к 30S субчастицам рибосом, что должно стимулировать инициацию трансляции.

Полиаденилирование мРНК в митохондриях и хлоропластах. Концепция *эндосимбиотического* происхождения митохондрий и хлоропластов

эукариот из бактерий-эндосимбионтов подкреплена многочисленными экспериментальными данными и является в настоящее время весьма популярной. В этой связи целесообразно рассмотреть особенности полиаденилирования РНК у этих органелл в данном разделе. Полиаденилированные мРНК митохондрий по своей структуре аналогичны мРНК *E. coli* класса IV (см. рис. 1.10,а). Как и у бактерий, полиаденилирование митохондриальных РНК происходит вне зависимости от специфических регуляторных последовательностей, характерных для мРНК эукариот. Средний размер поли(А)-последовательностей мРНК митохондрий в клетках HeLa человека составляет 55 нт, а в клетках асцитных опухолей мышей – 35–55 нт, что соответствует длине поли(А)-последовательностей у бактерий. Поли(А)-полимераза митохондрий клеток гепатомы Морриса обладает молекулярной массой 60 кДа. Она способна добавлять к РНК *in vitro* поли(А)-последовательности длиной до 600 нт, однако в изолированных митохондриях их размер составляет 20–23 нт. Фермент кодируется ядерным геном.

Длина 3'-концевых поли(А)-последовательностей РНК хлоропластов значительно превышает таковую РНК бактерий и достигает нескольких сотен нуклеотидов, что характерно для эукариотических поли(А)⁺-РНК. Эти последовательности не обязательно являются гомополимерами остатков аденозина, но могут состоять из кластеров А (75%), перемежающихся последовательностями G (24%), а также C и U (суммарное содержание – 5%), что напоминает свойства не кодирующих, обогащенных поли(А) участков мРНК бактериофага T7. Поли(А)-полимераза хлоропластов гороха (*Pisum sativum*) состоит из трех субъединиц, среди которых полипептид с молекулярной массой 43 кДа обладает антигенными детерминантами, общими с поли(А)-полимеразой дрожжей, а из двух других гликозилированных субъединиц лишь РНК-связывающий полипептид с молекулярной массой 105 кДа абсолютно необходим для функционирования фермента.

2.2.2. Редактирование пре-мРНК

Недавно появились сообщения о новых механизмах изменения кодирующего потенциала мРНК на посттранскрипционном уровне, названных *редактированием РНК* (editing). Оказалось, что в клетках многих организмов имеются ферментные системы, способные с высокой специфичностью

изменять первичную структуру мРНК, что, в свою очередь, меняет их кодирующий потенциал и приводит к образованию новых функционально значимых белков. Одним из первых был описан механизм редактирования РНК для митохондрий внутриклеточных паразитов – жгутиковых трипаносомид: *Crithidia fasciculata*, *Leishmania tarentolae* и *Trypanosoma brucei*. Митохондриальная (мт) ДНК этих одноклеточных представлена 20–50 идентичными копиями катенов кольцевых молекул (зацепленные друг за друга кольца в виде гирлянды), называемых *максикольцами*, которые являются функциональными аналогами мтДНК других эукариот. Кроме того, их митохондрии содержат несколько тысяч копий небольших молекул ДНК (*миникольца*), функциональное значение которых до недавнего времени оставалось неизвестным. В результате функционирования механизма редактирования мРНК в специфические участки митохондриальной мРНК встраиваются многочисленные остатки уридина (U), не кодируемые максикольцами митохондриального генома, тогда как другие остатки U, включенные в мРНК в результате транскрипции мтДНК, удаляются из транскриптов. При таком "редактировании" кодирующего потенциала мРНК могут быть изменены до нескольких сотен остатков U (табл. I.7), что приводит к образованию протяженных открытых рамок считывания (ОРС), кодирующих у жгутиковых высокомолекулярные белки, а также новых кодонов инициации трансляции (AUG), которые отсутствуют у мРНК, не подвергнутых редактированию. Такие модификации мРНК могут создавать новые терминирующие кодоны (UAG и UAA), а также затрагивать 3'-концевые нетранслируемые последовательности мРНК и poly(A)-хвосты. Функциональное значение редактирования этих нетранслируемых последовательностей мРНК неизвестно. Они, как полагают, могут оказывать влияние на стабильность соответствующих мРНК и их содержание внутри клеток.

Выбор последовательностей нуклеотидов в мРНК, которые подвергаются редактированию у трипаносомид, по-видимому, осуществляется с участием небольших *РНК-проводников* (guide RNAs, gRNA), частично комплементарных редактируемым участкам мРНК и кодируемых миникольцами мтДНК. gRNA являются первичными транскриптами, так как содержат на 5'-концах нуклеозиддифосфаты или нуклеозидтрифосфаты. На их 3'-концах обнаруживают 15 неcodируемых остатков U, которые добавляются

посттранскрипционно 3'-концевой уридилилтрансферазой.

Таблица I.7

Различные способы редактирования мРНК

Объект	Модифицированные или добавленные нуклеотиды
Митохондрии трипаносом Митохондрии <i>P. polycerphalum</i>	AAUUUAUGUUGUCUUU AUCUCUAAGGGUUUAACCGG
Парамиксовирусы (ген Р)	Сдвиг рамки AUUAAAAAGGGGCACAC
Митохондрии позвоночных	Сдвиг рамки CAGUAAAAAAAAAAAAA
высших растений	СТОП-кодон UUU <u>UUC</u> AUUGUGGUU <u>UAC</u>
Хлоропласты высших растений	Phe Tyr AAUAAU <u>U</u> GGCGAAACAU
Ионные каналы млекопитающих, регулируемые глутаматом	Иницирующий кодон UUUAUG <u>CGG</u> CAAGGA
Ген аполипопротеина В млекопитающих	Arg GAUAUAAUUUGAUCAGUAUA
	СТОП-кодон

Шесть или более 5'-концевых нуклеотидов gРНК комплементарны последовательности редактируемой мРНК, непосредственно предшествующей блоку нуклеотидов, подвергаемых редактированию. Уже сейчас ясно, что gРНК не служат матрицей для встраивания модифицируемых нуклеотидов в процессе посттранскрипционного редактирования РНК. Обсуждается гипотетический механизм этого процесса, который, к сожалению, не объясняет полностью имеющиеся факты. В соответствии с гипотезой образующиеся ошибочно спаренные (некомплементарные) нуклеотиды в гибридах gРНК и редактируемой мРНК служат сигналами для расщепления мРНК с последующим добавлением по местам разрывов остатков U. При этом места комплементарного спаривания нуклеотидов в РНК–РНК-гибридах защищены от редактирования. Кроме того, предполагают, что редактирование мРНК

происходит с участием сложных нуклеопротеидных комплексов – *эдитосом*, по аналогии со сплайсингом, где участие аналогичных комплексов – сплайсом в настоящее время доказано.

Другой тип редактирования РНК характерен для митохондрий слизевиков *P. polycephalum* (см. табл. I.7). В этом случае отдельные остатки С встраиваются во множественные участки митохондриальных мРНК, что приводит к сдвигам рамок считывания. Механизм данного процесса неизвестен.

Своеобразно решают задачу изменения кодирующего потенциала своих мРНК парамиксовирусы, в частности вирусы кори и свинки. Во время транскрипции гена Р, кодирующего белок, ассоциированный с полимеразой, РНК-полимераза совершает ошибки, что сопровождается вставками лишних остатков гуанозина (G) в мРНК и, как следствие, сдвигом рамок считывания при их трансляции.

В транскриптах митохондрий позвоночных животных в процессе полиаденилирования мРНК происходит создание бессмысленных кодонов UAA и UGA, приводящих к преждевременной терминации трансляции, что в конечном счете сопровождается появлением новых полипептидных цепей. Тот же самый механизм реализуется в ядрах позвоночных.

Другая группа механизмов редактирования мРНК основана на ферментативном взаимопревращении остатков нуклеотидов. Например, в большинстве митохондрий высших растений в результате дезаминирования происходит превращение остатков С в U. В меньшей степени для них характерен обратный процесс: $U \rightarrow C$. При этом изменяется смысл кодонов в мРНК и, как следствие, происходит замена соответствующих аминокислот в белках. В настоящее время еще не охарактеризованы ферментные системы, осуществляющие такие посттранскрипционные превращения нуклеотидов на уровне мРНК.

Отдельно следует упомянуть редактирование мРНК, происходящее в хлоропластах высших растений, в частности у кукурузы и табака. Оказалось, что предсказание последовательностей аминокислот в белках хлоропластов, сделанное на основании последовательностей нуклеотидов их генов, часто не соответствует действительности. Так, в гене *rp12* хлоропластов кукурузы и гене *psbL* хлоропластов табака находится кодон ACG в том месте, где ожидается расположение наиболее распространенного кодона инициации трансляции

АТГ. При созревании транскриптов этих генов происходит их посттранскрипционная модификация, сопровождаемая превращением С→U.

Своеобразный механизм посттранскрипционного редактирования РНК описан для транскриптов генов, кодирующих ионные каналы мозга млекопитающих, которые участвуют в передаче сигналов в синапсах центральной нервной системы. Известно, что субъединицы двух близкородственных классов глутаматных рецепторов содержат в определенных сегментах своих полипептидных цепей, формирующих каналы, остатки глутамина или аргинина, что оказывает существенное влияние на функционирование этих каналов. Оказалось, что субъединицы обоих классов кодируются генами, у которых в соответствующем месте имеется только кодон для глутамина (CAG), хотя в кДНК, полученной с использованием мРНК указанных субъединиц в качестве матрицы, в этом месте обнаружен также и аргининовый кодон CGG.

Интересно, что мРНК одной из субъединиц рецептора, принадлежащей на основании аминокислотных последовательностей к одному определенному классу (называемому здесь класс 1), подвергаются редактированию на 100%, тогда как транскрипты трех других субъединиц того же класса вообще не изменяются, несмотря на 90–95%-ную гомологию 30-звенных последовательностей, окружающих редактируемый сайт. Транскрипты двух субъединиц, принадлежащих к другому классу, редактируются с эффективностью соответственно 40 и 80%. Таким образом, во всех этих случаях имеет место количественная регуляция эффективности редактирования мРНК рецепторов мозга, что необходимо для пропорционального внутриклеточного синтеза соответствующих субъединиц. Механизм такой своеобразной регуляции экспрессии генов в настоящее время, к сожалению, не известен. Предполагают, что эдитосомы могут узнавать характерные элементы вторичной структуры мРНК, по-разному подвергающиеся редактированию.

Широкая, может быть, даже более широкая, чем мы сейчас представляем себе, распространенность посттранскрипционного редактирования мРНК у живых организмов указывает на этот механизм как на один из обычных способов реализации генетической информации. Использование таких модификаций пре-мРНК функционирующими

генетическими системами при экспрессии генов расширяет кодирующий потенциал генома и добавляет еще одну возможность регуляции их функционирования на посттранскрипционном уровне. В табл. I.8 суммированы данные об использовании редактирования мРНК у животных и некоторых их вирусов, которые наглядно демонстрируют распространенность этого явления у данной группы организмов.

В заключение рассмотрим еще один наиболее хорошо изученный пример редактирования мРНК аполипопротеина В (АРОВ) млекопитающих, участвующего в транспорте холестерина и триглицеридов в крови. Уже давно у млекопитающих обнаружили две формы АРОВ, кодируемые одним и тем же геном и образующиеся в результате редактирования АРОВ-мРНК (рис. I.11,а). Уникальный ген *АРОВ* человека содержит 29 экзонов, и его общая длина составляет 43 т.п.о. Длина мРНК этого гена, кодирующей АроВ100 с молекулярной массой 512 кДа, составляет 14 тысяч оснований (т.о.). В середине самого большого экзона 26 имеется глутаминовый кодон САА, который в результате редактирования мРНК превращается в терминирующий кодон UAA. Некоторые из мРНК, подвергшихся редактированию, расщепляются по *криптическому* (не функционирующему до расщепления) сайту полиаденилирования, что приводит к образованию более короткой

Таблица I.8

Редактирование РНК у животных и их вирусов

Организм, ткань	Локализация	РНК-субстрат	Последствия редактирования
Печень/кишечник крыс	Ядро	АРОВ-мРНК	C→U (CAA ^{Glu} →UAA ^{STOP})
Мозг человека и грызунов	»	мРНК рецепторов AMPA и КА	A→I (CAG ^{Glu} →CGG ^{Arg})
Семенники человека, нормальные и опухолевые ткани крыс	»	мРНК опухоли Вилмса 1	U→A (CUC ^{Leu} →CCC ^{Pro})
Мышцы человека	?	мРНК α-галактозидазы	U→A (TTC ^{Phe} →TAC ^{Tyr})
Опухоли человека	Ядро	мРНК нейрофиброматоза типа 1	C→U (CGA ^{Arg} →UGA ^{STOP})
Печень крыс	»	тРНК ^{Asp}	C→U и U→C рядом с антикодоновой петлей
Сумчатые	Митохондрии	тРНК ^{Gly}	C→U в антикодоне (Gly→Asp)
Вирус гепатита δ	Вирус	Антигеномная РНК	A→I (UAG ^{STOP} →UGG ^{Trp})
Парамиксовирусы	»	Ген Р	Вставки G
Вирус Эбола	»	Ген гликопротеина	Вставки A

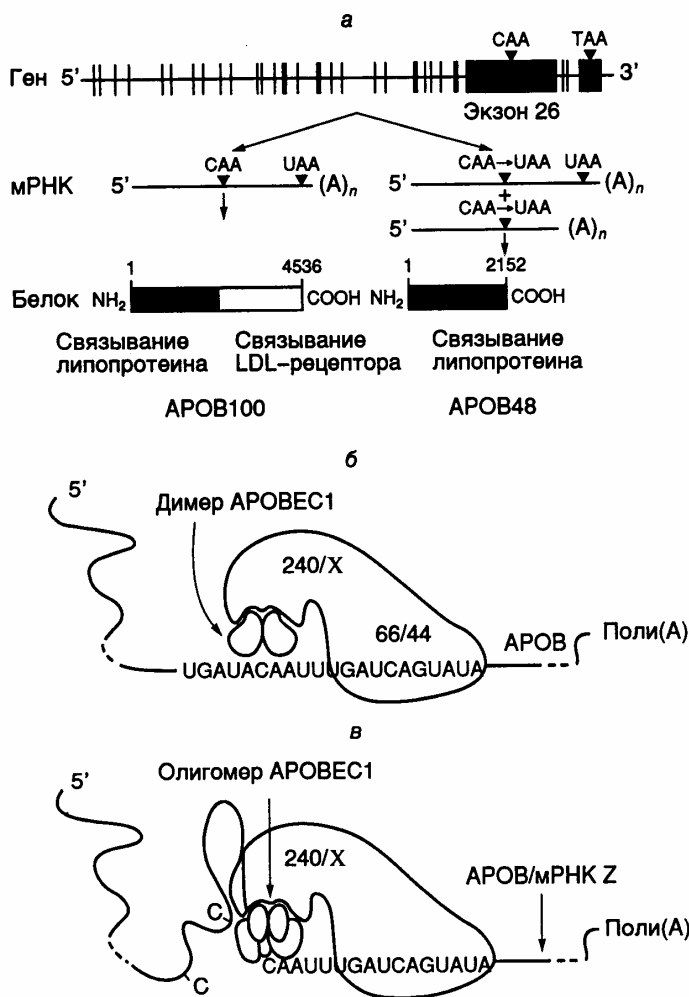


Рис. 1.11. Экспрессия гена *APOB* человека и механизм редактирования его мРНК

а – схема экспрессии гена; **б** – механизм дезаминирования остатка С в АРОВ-мРНК в результате редактирования с участием эдитосомы; **в** – нарушение специфичности редактирования АРОВ-мРНК при олигомеризации каталитической субъединицы АРОВЕС1. В этих условиях субстратом для эдитосомы могут быть другие мРНК (мРНК Z)

мРНК длиной в семь т.о. В результате трансляции отредактированных мРНК образуется укороченный АРОВ48 (241 кДа), который содержит 2152 N-концевых аминокислотных остатка АРОВ100. В функциональном отношении этот белок остается полностью активным, однако у него отсутствует С-концевой домен АРОВ100, который отвечает за связывание с рецептором липопротеинов низкой плотности. Молекулярным механизмом редактирования АРОВ-мРНК

является сайт-специфическое дезаминирование цитозина с помощью цитидиндезаминазы. Недавно клонировали ген, кодирующий этот фермент у человека, и полностью определили его первичную структуру. Показано, что цитидиндезаминаза человека, осуществляющая редактирование АРОВ-мРНК, представляет собой димер, построенный из двух идентичных субъединиц (АРОВЕС1) с молекулярной массой 28 кДа. Ген фермента, расположенный на хромосоме 12, экспрессируется исключительно в тонком кишечнике. Специфичность дезаминирования остатка цитозина определяется, по крайней мере, двумя факторами: последовательностью нуклеотидов мРНК в окрестностях этого сайта и белковыми кофакторами, взаимодействующими с каталитической субъединицей цитидиндезаминазы. Перенос методами генной инженерии последовательности нуклеотидов, фланкирующей сайт редактирования, в новые мРНК приводил к тому, что новые рекомбинантные РНК также подвергались специфическому редактированию *in vivo* и *in vitro*. В экспериментах такого рода, а также путем замен отдельных нуклеотидов методами направленного мутагенеза определили последовательность нуклеотидов мРНК, отвечающую за специфичность редактирования. Оказалось, что для оптимального осуществления этого процесса необходима последовательность длиной в 55 нуклеотидов, однако и 25 нуклеотидов в окрестностях сайта было достаточно, чтобы редактирование осуществлялось *in vitro* с 25%-ной эффективностью. Ниже представлены последовательности нуклеотидов в окрестностях сайтов редактирования, которые оказались высокоомологичными у разных видов млекопитающих.

			C→U		* * * * *
Человек	Сайт 1	GAUA	C	AAUU	<u>UGAUCAGUAUAUUAAG</u>
	Сайт 2	aAaA	C	AAUcca	UGAUCuacAUuUguuuu
Бабуин	Сайт 1	GAUA	C	AAUU	UGAUCAGUAUAUUAAG
Свинья	Сайт 1	GAUA	C	AAUU	UGAUCAGUAUAUUAAG
Кролик	Сайт 1	GAUA	C	AAUU	UGAUCAGUAUAUUAAG
Крыса	Сайт 1	GAUA	C	AAUU	UGAUCAGUAUAUUAagAG
Мышь	Сайт 1	GAUA	C	AAUU	UGAUCAGUAUAUUAAG

Звездочками обозначены нуклеотиды, замены которых с помощью направленного мутагенеза наиболее резко снижали эффективность редактирования, а строчными буквами отмечены негомологичные нуклеотиды. Подчеркнута последовательность нуклеотидов, получившая название *якорной последовательности* (mooring sequence). Эта последовательность является единственным *цис*-действующим регуляторным элементом, присутствие которого необходимо и достаточно для специфического редактирования вышерасположенного остатка C в экспериментах *in vitro*. Редактирование АРОВ-мРНК здесь происходит после удаления из нее интронов в результате сплайсинга. Эффективность редактирования в таких системах зависит от нуклеотидных последовательностей в окрестностях этого сайта. В частности, АТ-богатые последовательности без выраженной вторичной структуры, фланкирующие якорную последовательность, стимулируют редактирование АРОВ-мРНК.

В связи с тем, что простая 11-звенная якорная последовательность определяет нуклеотид, который редактируется в этой системе, возникает вопрос о дополнительных факторах, которые требуются для обеспечения специфичности функционирования системы редактирования, зависимой от присутствия якорной последовательности. Действительно, перемещение якорной последовательности *in vitro* к любому остатку C на расстоянии трех–четырех нуклеотидов от него в направлении 3'-конца РНК обеспечивает его редактирование *in vivo*. Предполагается, что в эдитосомах, осуществляющих редактирование АРОВ-мРНК, соблюдается определенное соотношение между содержанием АРОВЕС1-субъединиц и дополнительных факторов-помощников (см. рис. I.11,б). В этой части рисунка показана гипотетическая структура нормальной эдитосомы, осуществляющей редактирование АРОВ-мРНК. Такая эдитосома включает собственно каталитический димер, два полипептида с

молекулярными массами 66 и 44 кДа, необходимыми для специфического взаимодействия эдитосомы с якорной последовательностью, а также сопутствующие факторы X, ассоциированные с белком AUX240 (240 кДа), который регулирует эффективность редактирования мРНК, обеспечивая сборку эдитосомы из белковых компонентов.

Сверхэкспрессия АРОВЕС1-субъединицы в клетках, достигаемая генно-инженерными методами, приводит к изменению специфичности редактирования. Это было объяснено нарушением стехиометрических соотношений между молекулами субъединиц и дополнительными факторами, обеспечивающими специфичность редактирования (см. рис. 1.11, в). На рисунке показано, что избыток субъединицы АРОВЕС1 в эдитосоме стимулирует ее к редактированию вышерасположенных остатков С в АРОВ-мРНК, поскольку изменяется характер фолдинга 5'-концевой части редактируемой мРНК. В результате другие остатки С становятся доступными для каталитических субъединиц.

В соответствии с вышеизложенным, все известные в настоящее время формы редактирования пре-мРНК можно подразделить на два класса: *инсерционное редактирование* и *редактирование с замещением*. В первом случае редактирование мРНК сопровождается вставкой и/или удалением специфических нуклеотидов. Предполагается, что при такой форме редактирования последовательность нуклеотидов мРНК образует гибрид с gРНК, что сопровождается появлением неспаренных и ошибочно спаренных оснований, которые и маркируют сайты редактирования. Далее происходит расщепление углевод-фосфатного остова мРНК по этим сайтам в результате реакции трансэтерификации между фосфатными группами мРНК и встраиваемых нуклеотидов с последующим повторным лигированием образовавшихся фрагментов мРНК, сопровождаемым вставкой или удалением нуклеотидов. Такой механизм реализуется, по крайней мере, в митохондриях некоторых жгутиковых, а также слизевиков. У последних в результате редактирования митохондриальных мРНК имеют место вставки остатков С. Инсерционное редактирование мРНК парамиксовирусов, сопровождающееся вставками остатков G, по-видимому, происходит вследствие ошибок РНК-полимеразы при транскрипции соответствующих генов. В результате редактирования с замещением, как это имеет место в случае редактирования

АРОВ-мРНК, а также, возможно, мРНК митохондрий и хлоропластов высших растений и ионных каналов, не происходит расщепления фосфодизэфирных связей в редактируемой РНК, а новое азотистое основание синтезируется *in situ* (т.е. модифицируется непосредственно в ее полинуклеотидной цепи). При этом не используется gРНК. Формально такой механизм напоминает реакции посттранскрипционной модификации азотистых оснований в тРНК, рРНК, малых ядерных РНК, а также двухцепочечных РНК.

Каково же биологическое значение механизма редактирования генетической информации на уровне мРНК? Какие силы заставили этот механизм эволюционно закрепиться у большого числа далеко отстоящих друг от друга биологических видов? Почему для организмов выгоднее изменять информацию посттранскрипционно, а не заключать ее непосредственно в генах? Очевидно, что для этого должны быть веские, не вполне понятные сегодня причины, которые не допускают перехода к обычному кодированию такой информации. На мой взгляд, редактирование мРНК может иметь непосредственное отношение к дополнительной стабилизации генетической информации в наиболее уязвимых для мутагенеза генетических локусах. Действительно, редактирование РНК получило наибольшее распространение в хлоропластах и митохондриях высших организмов, а также у одноклеточных эукариот. В разделе 5.3.1 приводится обоснование того, что дополнительная защита генетической информации от разрушительного действия химических мутагенов особенно нужна именно многоклеточным организмам для предотвращения накопления генетического груза в делящихся соматических клетках при онтогенезе. Вероятно, одним из путей достижения этого было эволюционное включение в геном эукариот избыточных последовательностей нуклеотидов. У свободноживущих одноклеточных организмов *на популяционном уровне* такой проблемы не существует, поскольку гибель отдельной свободноживущей клетки не грозит существованию популяции этих клеток, как это имеет место у Metazoa. Однако у митохондрий и хлоропластов, часто рассматриваемых в качестве внутриклеточных микроорганизмов-эндосимбионтов, наблюдаются совершенно особые условия существования. Несмотря на то что их геном содержит мало избыточных последовательностей и, следовательно, слабо защищен ими от химических мутагенов, мутации в определенных генах митохондрий и хлоропластов могут быть летальными для

соматической клетки-хозяина и всего многоклеточного организма. В этих условиях мутационное изменение нуклеотидов, подвергающихся редактированию на уровне РНК, фактически заменяет само редактирование, и такие мутации нейтральны в функциональном отношении. В отсутствие мутаций редактирование корректирует первичную структуру РНК, а при наличии их необходимость в редактировании отпадает. Иными словами, во всех этих локусах редактирование как бы упреждает мутационные замены редактируемых нуклеотидов в генах, которые геном по каким-то причинам не может эффективно предотвратить в силу особенностей структуры и функционирования соответствующих генетических локусов. По аналогии с неоднозначностью генетического кода наличие механизма редактирования допускает сосуществование в конкретных генетических локусах "вырожденных сайтов", различающихся по первичной структуре, но не своему генетическому смыслу.

2.2.3. Другие модификации эукариотических мРНК

Посттранскрипционные модификации предшественников эукариотических мРНК по сравнению с теми же изменениями первичных транскриптов прокариот более разнообразны и играют большую роль в регуляции экспрессии их генов. Почти все эти реакции происходят в ядре эукариотических клеток в процессе синтеза РНК или сразу же после его завершения. Прежде всего, к 5'-концевому нуклеотиду большинства пре-мРНК присоединяются *кэп-группы*, что сопровождается метилированием одного или нескольких концевых нуклеотидов, в большинстве случаев необходимым для стабилизации и экспорта соответствующих мРНК из ядра в цитоплазму, а также их эффективной трансляции рибосомами. В основном те же функции, по-видимому, выполняет и полиаденилирование 3'-концевых последовательностей мРНК, которые подготавливаются к этому этапу путем специфического отщепления избыточных 3'-концевых нуклеотидов предшественника. Кроме того, интроны, содержащиеся в гигантских первичных предшественниках мРНК, с высокой точностью удаляются в результате сплайсинга. Ниже будут рассмотрены механизмы перечисленных посттранскрипционных модификаций пре-мРНК эукариот. Об использовании этих реакций в регуляции экспрессии

генов на посттранскрипционном уровне речь пойдет в разделе 3.3.

Кэпирование. Сразу же после инициации транскрипции чаще всего происходит котранскрипционная модификация 5'-конца мРНК, сопровождаемая присоединением так называемой кэп-группы и дальнейшими ее изменениями. Кэпирование является одной из самых ранних модификаций растущих цепей РНК и происходит после полимеризации ее первых 20–30 нуклеотидов. Такая котранскрипционная модификация мРНК не только стабилизирует мРНК в

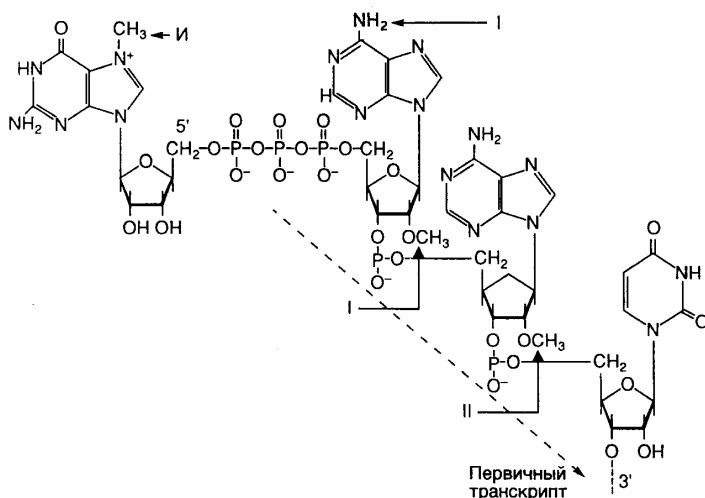


Рис. 1.12. Обобщенная структура кэп-группы эукариотических мРНК

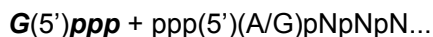
Указаны сайты метилирования, характерные для кэп-групп разных типов.

И – инвариантный сайт, метилируемый у кэп-групп всех типов, I, II – сайты, метилируемые у кэп-групп типов 1 и 2 соответственно

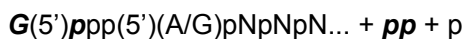
цитоплазме, но и необходима в большинстве случаев для ее эффективной трансляции. Так, один из факторов инициации трансляции eIF-4E выполняет функции кэп-связывающего белка и требуется для осуществления кэп-зависимой трансляции мРНК. Кроме того, установлено, что кэпирование мРНК необходимо для эффективного сплайсинга пре-мРНК, ее полиаденилирования и экспорта из ядра в цитоплазму. Кэпированию подвергаются только транскрипты РНК-полимеразы II. На исключительную значимость реакций кэпирования указывает и тот факт, что контролирующие их гены являются жизненно важными.

Как уже упоминалось выше, транскрипция у эукариот и прокариот начинается, как правило, с пуринового рибонуклеозидтрифосфата – АТР, или GTP, причем трифосфатная группа сохраняется в составе мРНК. Таким

образом, 5'-концевая последовательность мРНК в ядре на ранних этапах транскрипции представлена в следующем виде: $ppp(A/G)pNpNpN...$ Гуанилилтрансфераза катализирует присоединение к растущей цепи мРНК молекулы GMP, которая оказывается связанной с 5'-концевым пурином 5'-5'-трифосфатной группой. Суммарная реакция первого этапа процесса кэпирования выглядит следующим образом:



↓ Гуанилилтрансфераза



Реакция, по-видимому, протекает в две стадии. Вначале фермент связывает молекулу GTP (входящую затем в состав кэп-группы), что сопровождается отщеплением пирофосфата и образованием ковалентной связи фермент–GMP. Далее GMP присоединяется к 5'-концу мРНК, которая в результате теряет γ -фосфатную группу. В результате нуклеотид кэп-группы оказывается в обратной ориентации по отношению к остальным нуклеотидам мРНК. Процесс создания кэп-группы этой последовательностью реакций не заканчивается. На заключительных этапах кэпирования происходит метилирование по N7 ранее присоединенной молекулы гуанозина. Такие посттранскрипционные модификации происходят в несколько стадий в цитоплазме клеток после транспорта процессированной мРНК из ядра с участием цитоплазматических ферментов.

Первая стадия метилирования осуществляется ферментом РНК(гуанил-7)-метилтрансферазой, которая переносит метильную группу S-аденозилметионина в положение 7 концевого гуанина кэп-группы (рис. 1.12). Кэп-группа, метилированная лишь по этому положению, характерна для одноклеточных эукариот и получила название *кэпа 0-го типа*. Вслед за этим у большинства многоклеточных эукариот происходит метилирование 2'-ОН рибозы 5'-концевого инициаторного нуклеотида (А или G), который является первым нуклеотидом, включаемым в мРНК при инициации ее синтеза РНК-полимеразой. Метилирование катализирует другой цитоплазматический

фермент – 2'-О-метилтрансфераза. Такая основная форма кэпа большинства эукариот получила название *кэпа 1-го типа*. Очень редко и только у тех мРНК, инициация синтеза которых происходит с АТР, под действием 2'-О-метиладенозин-N⁶-трансферазы метилируются NH₂-группы этого остатка А. Фермент распознает данную концевую группу в качестве субстрата лишь в том случае, если она была предварительно метилирована в положении 2'-ОН в результате вышеописанной реакции.

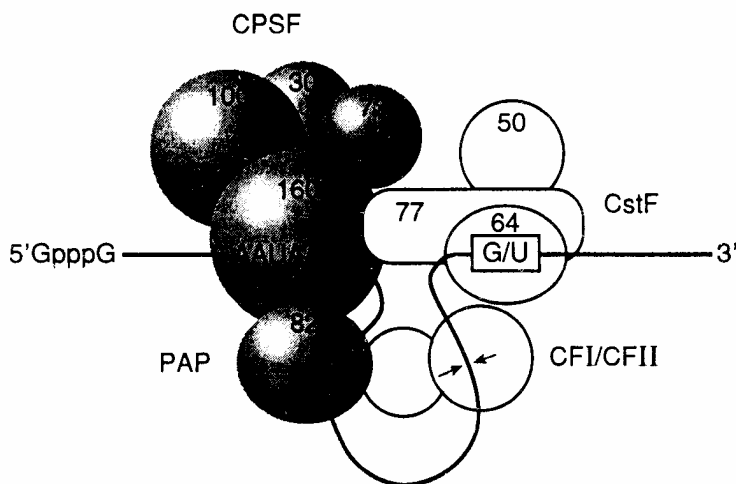


Рис. I.13. Модель процессинга 3'-концевых последовательностей эукариотических пре-мРНК

Цифрами указаны молекулярные массы белков, входящих в состав процессирующего комплекса. Обозначены кэп-группа, последовательность поли(А)-сайта и регуляторная последовательность, с которой взаимодействует гетеродимерный фактор, стимулирующий расщепление (CstF). Стрелки указывают место расщепления РНК

PAP – поли(А)-полимераза, CPSF – фактор специфичности и расщепления РНК, CF I и CF II – факторы расщепления I и II

У некоторых видов эукариот метильная группа может дополнительно присоединяться ко второму от кэп-нуклеотида нуклеозиду мРНК (см. рис. I.12). Субстратом для этого фермента служит мРНК с кэпом 1-го типа, уже содержащим две метильные группы. В результате происходит метилирование остатка рибозы по 2'-ОН-группе с образованием структуры, получившей название *кэпа 2-го типа*. Если эта реакция имеет место, то мРНК, содержащие кэп 2-го типа, составляют 10–15% от общей популяции молекул кэпированных мРНК.

Иная структура кэп-группы характерна для некоторых зрелых некодирующих РНК, в частности малых ядерных РНК, обогащенных урацилом (У-мяРНК). В этом случае остаток гуанозина кэп-группы дважды метилирован в положении 2 в дополнение к обычной метильной группе в положении 7: $m^{2,2,7}G(5')ppp(5')N$. Такое гиперметилирование У-мяРНК требуется для импорта собранных У-мяРНК-частиц в ядро и, возможно, предотвращает вовлечение У-мяРНК в трансляцию.

Полиаденилирование. Одним из обязательных этапов созревания предшественников эукариотических мРНК, синтезированных в ядре, является процессинг их 3'-концевых последовательностей, тесно сопряженный с присоединением кэп-группы. Созревание 3'-конца мРНК является двухэтапным процессом. Вначале предшественник теряет 3'-концевую некодирующую последовательность, после чего, как правило, к 3'-концу присоединяется поли(А)-последовательность путем ферментативной полимеризации остатков AMP:



↓ Расщепление



↓ Полиаденилирование



В настоящее время известно несколько исключений из этого правила: гистоновые мРНК животных и мРНК некоторых вирусов, предшественники которых расщепляются с помощью высокоспецифических эндонуклеаз и не полиаденилируются. Остаются неполиаденилированными и У-мяРНК, которые также являются транскриптами РНК-полимеразы II. В этом случае кэпированный первичный транскрипт мяРНК U1, содержащий на своем 3'-конце несколько избыточных нуклеотидов, экспортируется из ядра в цитоплазму, где и происходит удаление избыточной последовательности, которое в ядре заблокировано специфическим белковым ингибитором TPI (3'-terminal processing inhibitor).

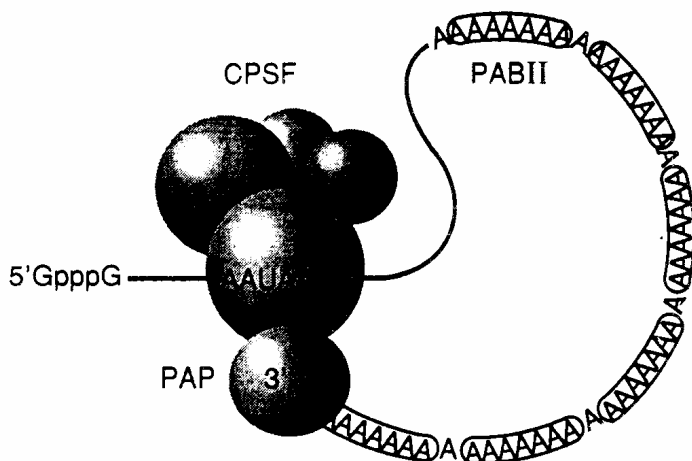


Рис. I.14. Модель белкового комплекса, элонгирующего поли(A)
 PAP – поли(A)-полимераза, PAB II – поли(A)-связывающий белок II

Места отщепления 3'-концевых некодирующих последовательностей в мРНК животных обычно маркированы специальными последовательностями нуклеотидов (рис. I.13). Имеются, по крайней мере, две такие последовательности, образующие *сайты полиаденилирования*, или *поли(A)-сайты*. Одна из них – AAUAAA расположена за 15 нуклеотидов перед расщепляемой фосфодиэфирной связью и практически одинакова у всех исследованных организмов. Другая, менее изученная последовательность располагается сразу же за первой и часто состоит из нескольких остатков U или обогащена GU. Сайт расщепления РНК определяется расстоянием между этими двумя элементами с предпочтительным расщеплением фосфодиэфирной связи на 3'-конце остатка А, находящегося на участке, в котором расщепление разрешено. Имеются данные о том, что последовательности, расположенные выше AAUAAA, могут оказывать стимулирующее влияние на процессинг, но их присутствие необязательно для его правильного осуществления.

С последовательностью AAUAAA взаимодействует фактор CPSF (cleavage and polyadenylation specificity factor), определяющий специфичность расщепления и полиаденилирования РНК. CPSF состоит из четырех субъединиц с молекулярными массами 160, 100, 70 и 30 кДа. Последняя из них, по-видимому, не является необходимой для его функционирования. Самая

большая субъединица находится в непосредственном контакте с последовательностью AAUAAA.

С GU-богатой и ниже расположенной последовательностью связывается гетеродимерный белковый фактор CSTF (cleavage stimulating factor), стимулирующий расщепление и состоящий из трех субъединиц (77, 64 и 50 кДа). Вторая субъединица контактирует с GU-богатым регуляторным элементом и обладает типичным РНК-связывающим доменом. По отдельности факторы CPSF и CSTF лишь слабо взаимодействуют с РНК. Однако их одновременное присутствие приводит к образованию прочного комплекса. Такой кооперативный эффект и взаимодействие двух факторов между собой определяются их большими субъединицами.

В расщеплении РНК непосредственно участвуют еще два фактора: CFI и CFII (cleavage factors). Как и в предыдущем случае, лишь вместе они образуют прочный комплекс с РНК.

Для полного реконструирования бесклеточной системы, осуществляющей процессинг 3'-концов *in vitro*, в ней помимо вышеупомянутых факторов необходимо наличие поли(А)-полимеразы – фермента, непосредственно осуществляющего полиаденилирование. Присутствие этого фермента требуется не для самого акта расщепления РНК, а, по-видимому, для стабилизации процессирующего белкового комплекса, схематически изображенного на рис. 1.13. Сборка такого сложного комплекса зависит от АТФ, однако в процессе сборки не происходит расщепления ее β - γ -связей. В настоящее время неизвестно, какой именно компонент этого комплекса непосредственно расщепляет фосфодиэфирные связи РНК. Процесс полиаденилирования начинается сразу же за расщеплением РНК и происходит настолько быстро, что неполиаденилированных промежуточных продуктов не обнаруживается. Такое сопряжение двух реакций необходимо для защиты 3'-концевых последовательностей РНК от деградации нуклеазами. При этом сам акт полиаденилирования требует наличия только фактора CPSF, но не трех других: CSTF, CFI и CFII.

Поли(А)-полимераза животных состоит из двух субъединиц с молекулярными массами ~ 80 и ~ 43 кДа, которые образуются в результате альтернативного сплайсинга их общей пре-мРНК. Короткий полипептид не обладает ферментативной активностью, и его функции неизвестны. Большая

полипептидная цепь содержит С-концевой домен, обогащенный Ser и Thr и не определяющий ни одну из функций фермента, обнаруживаемых *in vitro*. Предполагают, что регуляторную роль играет множественное фосфорилирование этого домена. С-Концевой домен также содержит один из двух сигнальных последовательностей, необходимых для транспорта фермента в ядро. Вторая сигнальная аминокислотная последовательность локализована на границе С-концевого домена и основного полипептида поли(А)-полимеразы. Сравнительное исследование первичной структуры поли(А)-полимеразы показало наличие в ее полипептидной цепи каталитического домена, характерного для полимераз так называемого *семейства X*, к которому относятся ДНК-полимераза β , терминальная трансфераза, а также некоторые другие нуклеотидилтрансферазы. Используя 3'-конец расщепленной РНК в качестве затравки, поли(А)-полимераза последовательно присоединяет к нему остатки АМР из АТР по тому же самому механизму, что и другие ДНК- и РНК-полимеразы. Для эффективного функционирования поли(А)-полимераза требует наличия фактора CPSF, а также поли(А)-связывающего белка PAB II (poly(A)-binding protein II), который связывает полиаденилирующий комплекс с РНК после присоединения к ней, по крайней мере, десяти остатков А. В присутствии этих двух факторов поли(А)-полимераза сразу синтезирует поли(А)-последовательность полной длины по процессивному механизму. Гипотетическая структура элонгирующего комплекса представлена на рис. I.14.

Процессивное (непрерывное) полиаденилирование 3'-концов РНК происходит со скоростью ~ 25 нуклеотидов/с до тех пор, пока длина поли(А)-последовательности не достигнет ~ 250 нуклеотидов. После этого процессивная реакция прекращается, и происходит медленное дистрибутивное присоединение остатков АМР разными молекулами поли(А)-полимеразы. Предполагают, что элонгирующий белковый комплекс узнает длину синтезированной поли(А)-последовательности при участии фактора PAB II (см. рис. I.14). По этому механизму связывание определенного числа молекул PAB II с поли(А) прекращает элонгацию поли(А)-последовательности. Такой строгий контроль за длиной поли(А) на 3'-концах процессированных мРНК имеет большое значение для действия механизма, контролирующего время

полужизни мРНК в цитоплазме. Без тщательного контроля над этим процессом с помощью селективного деаденилирования невозможно регулировать внутриклеточную деградацию мРНК, а вместе с тем и уровень экспрессии соответствующих генов с участием данного механизма.

Полиаденилирование является универсальным феноменом, играющим важную роль в процессинге и функционировании мРНК как прокариотических, так и эукариотических организмов. Однако сравнение механизмов полиаденилирования у этих групп организмов выявляет существенные различия, суммированные в табл. I.9.

Н. Саркаром (1997 г.) было высказано предположение о возникновении механизма полиаденилирования РНК. Он полагает, что поскольку у бактериальных и эукариотических ферментов, выполняющих аналогичные функции, не обнаружено гомологии в аминокислотных последовательностях, оба фермента возникли недавно из уже значительно дивергировавших функционально родственных предшественников. Такими предшественниками могли быть прокариотические и эукариотические тРНК-нуклеотидилтрансферазы, осуществляющие посттранскрипционный синтез последовательности ССА на 3'-концах тРНК, что по своему механизму близко к полиаденилированию. Подтверждением этого является обнаруженная значительная гомология между тРНК-нуклеотидилтрансферазой и основной поли(А)-полимеразой *E. coli*, а также между тРНК-нуклеотидилтрансферазой бактерии *Sulfolobus shibatae*, обитающей в горячих серных источниках, и поли(А)-полимеразами эукариот. Следы такой эволюционной связи обнаруживаются и в современных митохондриях, где в результате полиаденилирования мРНК могут создаваться терминирующие кодоны. По мнению Саркара, различия между системами полиаденилирования прокариот и эукариот, представленные в табл. I.9, можно рассматривать в качестве продукта эволюционной дивергенции сходных биосинтетических функций в процессе независимого возникновения нового регуляторного механизма, обеспечивающего физиологические нужды уже глубоко различающихся групп организмов.

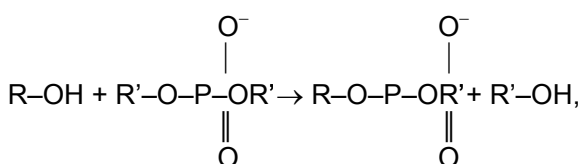
Таблица I.9

Сравнение полиаденилирования мРНК у эукариот и прокариот

Функции	Млекопитающие	E. coli
Длина поли(А)-последовательностей, нт	80–200	14–60
Уровень полиаденилирования отдельных мРНК, %	~100	2–50
Локализация сайтов полиаденилирования	Ниже консенсусной последовательности AAUAAA	Любые доступные 3'-ОН-концы мРНК
Функции:		
стабильность мРНК	Стабилизация мРНК без участия 3'-концевых структур типа "стебель–петля"	Дестабилизация мРНК, обладающих 3'-концевыми структурами типа "стебель–петля", возможная стабилизация мРНК без таких структур
трансляция	Специфический контакт с белком PABP, необходимый для взаимодействия мРНК с 40S субчастицами рибосом	Связывание рибосомного белка S1, возможно, необходимое для взаимодействия мРНК с 30S субчастицами рибосом
репликация плазмид	Не участвует	Деградация антисмысловой РНК, ингибирующей репликацию плазмид типа ColE1

(Рy)_n перед 3'-концевым сайтом сплайсинга существенна для правильного вырезания интронов. Остаток аденозина в консервативной последовательности нуклеотидов интрона, расположенный ближе к его 3'-концу, получил название *точки разветвления* (branch point). Именно с этим аденозином ковалентно соединяется 5'-конец интрона, освобождающийся на первом этапе сплайсинга с образованием *структуры типа "лассо"* (lariat) (см. ниже). Первичная структура указанных сайтов мало консервативна в генах, кодирующих ядерные пре-мРНК, и может значительно варьировать даже у интронов одного и того же организма. В зависимости от механизма вырезания интронов и особенностей их пространственной структуры различают *интроны групп I, II и III*, интроны ядерных РНК, а также *твинтроны* – интроны, расположенные внутри интронов.

Аутосплайсинг интронов групп I, II и III. Открытие интронов, удаляющихся из пре-РНК в результате аутокаталитического процесса – аутосплайсинга, имело далеко идущие последствия для развития фундаментальных и прикладных исследований в молекулярной генетике и породило целое направление с использованием *рибозимов* (подробнее см. раздел 9.2). Такие интроны впервые были обнаружены в органеллах высших организмов, но вскоре было продемонстрировано их широкое распространение в природе. Исследование молекулярных механизмов аутосплайсинга показало, что хотя все вышеупомянутые интроны в сплайсинге используют одни и те же реакции *трансэтерификации*:



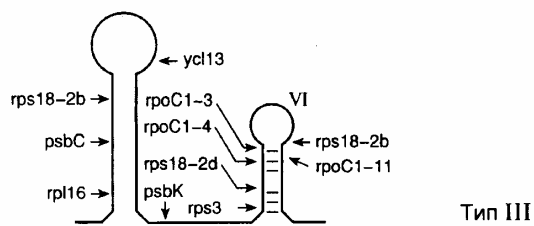
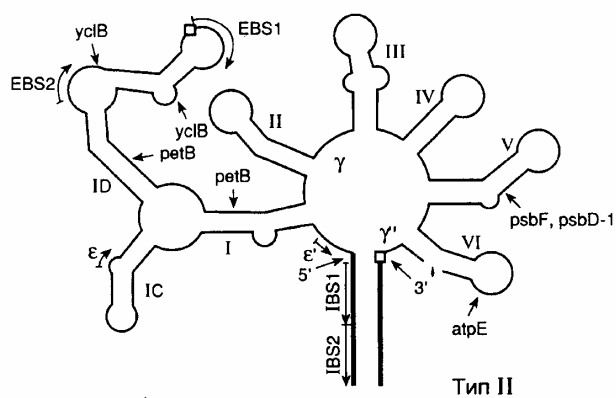
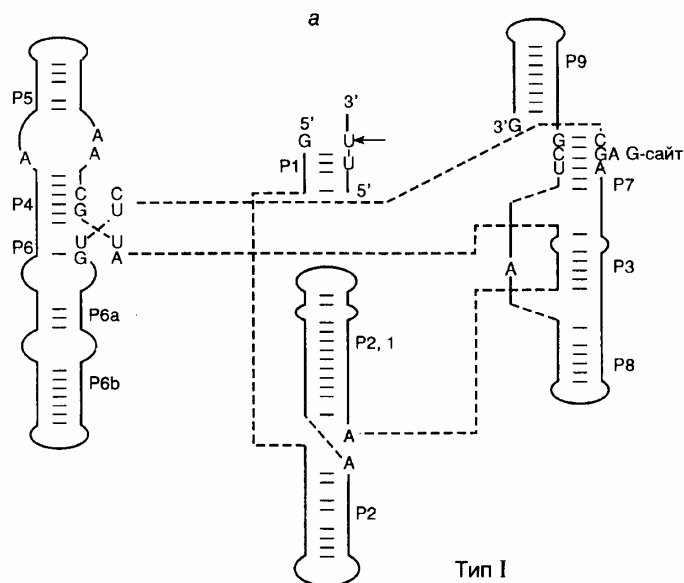
детали механизмов для интронов групп I, II и III различны.

На рис. I.15,а,б изображены структурные особенности интронов I–III групп. Интроны группы I образуют наиболее сложные вторичные и третичные структуры. Они обнаружены в предшественниках РНК простейшего *Tetrahymena thermophila* и удаляются из них с использованием приведенного в табл. I.10 механизма.

Таблица I.10

**Механизм прямой и обратной реакций аутосплайсинга
интронов группы I**

Аутосплайсинг	Обратное лигирование
GTP_{OH} $+$ $[\text{Экзон } 1]\text{urA}-\text{Инtron}-\text{GpA}[\text{Экзон } 2]$ $\downarrow$ Расщепление 5'-концевого сайта сплайсинга $\downarrow$ $[\text{Экзон } 1]\text{u}-\text{OH}$ $\text{GpA}-\text{Инtron}-\text{GpA}[\text{Экзон } 2]$ $\downarrow$ Лигирование экзонов $\downarrow$ $[\text{Экзон } 1]\text{ура}[\text{Экзон } 2]$ $+$ $\text{GpA}-\text{Инtron}-\text{G}$	GTP_{OH} $+$ $[\text{Экзон } 1]\text{urA}-\text{Инtron}-\text{GpA}[\text{Экзон } 2]$ $\uparrow$ Интеграция интрона в прежнее место $\uparrow$ $[\text{Экзон } 1]\text{u}-\text{OH}$ $\text{GpA}-\text{Инtron}-\text{GpA}[\text{Экзон } 2]$ $\uparrow$ Лигирование экзона 2 и интрона $\uparrow$ $[\text{Экзон } 1]\text{ура}[\text{Экзон } 2]$ $+$ $\text{GpA}-\text{Инtron}-\text{G}$



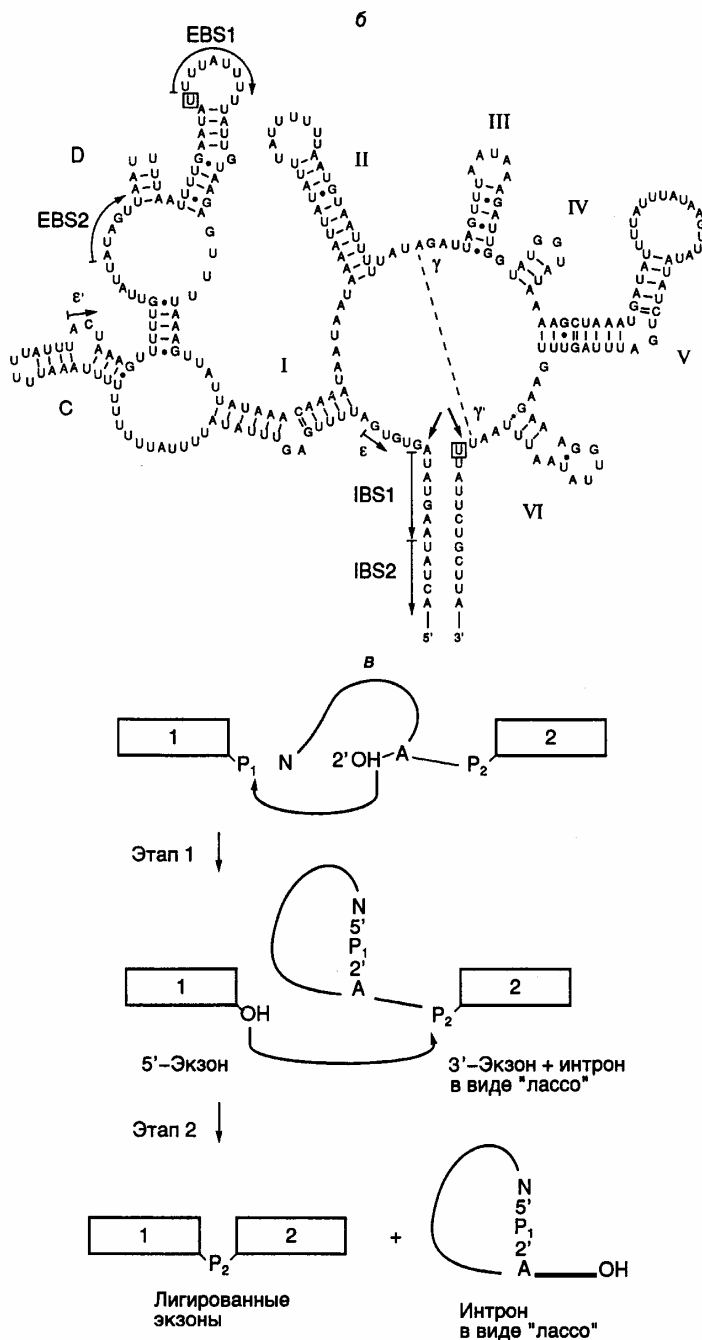


Рис. I.15. Структура интронов I-III групп и механизм аутосплайсинга
а – вторичные структуры типичных интронов типа I (*Tetrahymena thermophila*), II и III групп (*Euglena gracilis*). Стрелками обозначены 3'- и 5'-сайты сплайсинга, а также места интеграции твинтронов, пунктирными линиями – сайты интронов, сближенные в пространстве, EBS и IBS – соответственно сайты связывания экзонов и интронов; **б** – предполагаемая вторичная структура интрона группы II; **в** – двухэтапный

механизм аутосплайсинга. Римскими цифрами обозначены предполагаемые двухцепочечные участки РНК, 1, 2 – экзоны

Строчными буквами в таблице изображены нуклеотиды, принадлежащие экзонам, а прописными – интронам. Прямая и обратная реакции не требуют белковых катализаторов. Однако в ряде случаев интроны I и II групп могут кодировать полипептид, названный *матуразой* (maturase, от англ. mature – созреть), который необходим для эффективного аутосплайсинга и транспозиции интронов *in vivo*.

Аутосплайсинг интронов группы I начинается с нуклеофильной атаки 2'-ОН-группы GTP по фосфодиэфирной связи в 5'-концевом сайте сплайсинга (см. табл. I.10) (присутствие GTP в качестве кофактора необходимо для протекания данной реакции). При этом происходит освобождение 5'-концевого экзона 1. Образовавшиеся половинки молекулы удерживаются рядом друг с другом водородными связями. Освободившаяся 3'-ОН-группа экзона 1 далее атакует фосфодиэфирную связь в 3'-концевом сайте сплайсинга, что сопровождается расщеплением этой связи и полным освобождением интрона. Одновременно происходит лигирование двух экзонов. Все этапы этой реакции обратимы. В обратной реакции после связывания соответствующего участка лигированных экзонов "активным центром" вырезанного интрона последний атакует своей концевой 3'-ОН-группой фосфодиэфирную связь в месте стыковки экзонов, что сопровождается объединением интрона и экзона 2. На втором этапе 3'-ОН-группа экзона 1 атакует фосфатную группу по первой фосфодиэфирной связи интрона, освобождая остаток G, с последующим объединением экзона 1 с остальной частью молекулы.

Аутосплайсинг интронов II и III групп протекает по другому механизму (см. рис. I.15,в). Реакция также осуществляется в два этапа. На первом этапе 2'-ОН-группа аденозина, находящегося в точке разветвления интрона, вырезаемого из предшественника РНК, атакует фосфодиэфирную связь 5'-концевого сайта сплайсинга, что сопровождается освобождением экзона 1 и образованием промежуточной структуры в виде лассо. В таком промежуточном соединении 5'-концевой нуклеотид интрона соединен 2'-5'-фосфодиэфирной связью с этим остатком аденозина, образуя петлю на конце интрона. На втором этапе 3'-концевая ОН-группа экзона 1 атакует 3'-концевой сайт сплайсинга. В

результате происходит объединение экзонов и освобождается интрон с петлей на 5'-конце. В заключение следует отметить одно замечательное свойство этой группы реакций. Поскольку число фосфодиэфирных связей на протяжении всех преобразований остается неизменным, они не требуют затраты дополнительной энергии и протекают самопроизвольно.

Сплайсинг у ядерных пре-мРНК. Исследование механизмов сплайсинга у ядерных пре-мРНК показало, что они существенно не отличаются от только что рассмотренных механизмов, используемых для удаления интронов II и III групп, однако ряд различий имеется.

Структура интронов, удаляемых из предшественников РНК аутосплайсингом, высококонсервативна. У интронов ядерных РНК консервативны лишь короткие последовательности в окрестностях сайтов сплайсинга и точки разветвления. Вторая отличительная особенность сплайсинга у ядерных пре-мРНК – потребность в АТР и мяРНП, формирующих сплайсому. Тем не менее, одинаковые черты этих двух систем сплайсинга более существенны. Малые ядерные РНК, входящие в состав мяРНП и сплайсома, участвуют в распознавании сайтов сплайсинга и точки разветвления, а также в правильном ориентировании экзонов по отношению друг к другу. Их ОН-группы осуществляют атаку фосфодиэфирной связи 5'-концевого сайта сплайсинга на первом этапе реакции. Так же как и в случае интронов II и III групп, для вырезаемого интрона характерна структура типа лассо.

Распознавание канонических сайтов при сплайсинге ядерных пре-мРНК. Для полного прохождения всех реакций сплайсинга его аппарат (сплайсома) должен распознавать на пре-мРНК три критические последовательности нуклеотидов: 3'- и 5'-концевые сайты сплайсинга и точку разветвления. В соответствии с наиболее распространенной моделью поиска экзонов в пре-мРНК с большими интронами, аппарат сплайсинга, прежде всего, отыскивает два близкорасположенных сайта сплайсинга в нужной ориентации. Распознавание конкретной пары сайтов сплайсинга в пре-мРНК дрожжей начинается с взаимодействия по принципу комплементарности между последовательностями этих сайтов и 5'-концевыми участками мяРНК U1, входящих в состав U1-мяРНП-частиц в случае 5'-концевых сайтов сплайсинга, а также белка U2AF⁶⁵, находящегося в комплексе с U2AF³⁵, с

последовательностью 3'-концевого сайта сплайсинга. Эффективность функционирования сайтов сплайсинга зависит от степени соответствия их первичной структуры консенсусным последовательностям.

Сближение двух сайтов сплайсинга, сопровождаемое выпетливанием последовательности интрона, осуществляется с участием SR-белков, обогащенных Ser и Arg, которые взаимодействуют с 70K-белком (70 кДа), входящим в состав U1-мяРНК-частицы, а также U2AF³⁵ на другом конце интрона. Такие комплексы, образованные с участием SR-белков, получили название "*коммитированных комплексов*".

В то время как 5'-концевые сайты сплайсинга довольно консервативны, 3'-концевые сайты обнаруживают большую вариабельность. U1-мяРНК, находящиеся в комплексе с 5'-концевым сайтом сплайсинга, стимулируют присоединение белков U2AF к 3'-концевому сайту и способствуют правильному распознаванию соответствующего экзона в процессе, получившем название "*определение экзона*". Эффективное распознавание сайтов сплайсинга соответствующими белками является критическим этапом в правильном и упорядоченном удалении экзонов из пре-мРНК в процессе конститутивного сплайсинга.

Механизмы определения экзонов. Гены позвоночных, как правило, содержат большое число мелких экзонов (средняя длина ~ 137 п.о.), разделенных протяженными последовательностями интронов, длина которых может превышать 100 т.п.о. При этом ферментные системы, осуществляющие сплайсинг, должны распознавать небольшие экзоны, затерянные среди гигантских некодирующих последовательностей, и осуществлять их правильное объединение друг с другом. Как уже упоминалось выше, канонические (консенсусные) последовательности, необходимые для распознавания экзонов, располагаются на концах интронов. Однако в отличие от соответствующих последовательностей интронов дрожжей регуляторные последовательности в окрестностях сайтов сплайсинга животных менее консервативны, и это, на первый взгляд, значительно усложняет задачу распознавания экзонов системами сплайсинга. В действительности распознавание индивидуальных сайтов сплайсинга у животных, по-видимому, не связано с независимым распознаванием канонических последовательностей в окрестностях каждого экзона. В соответствии с наиболее популярной в настоящее время моделью

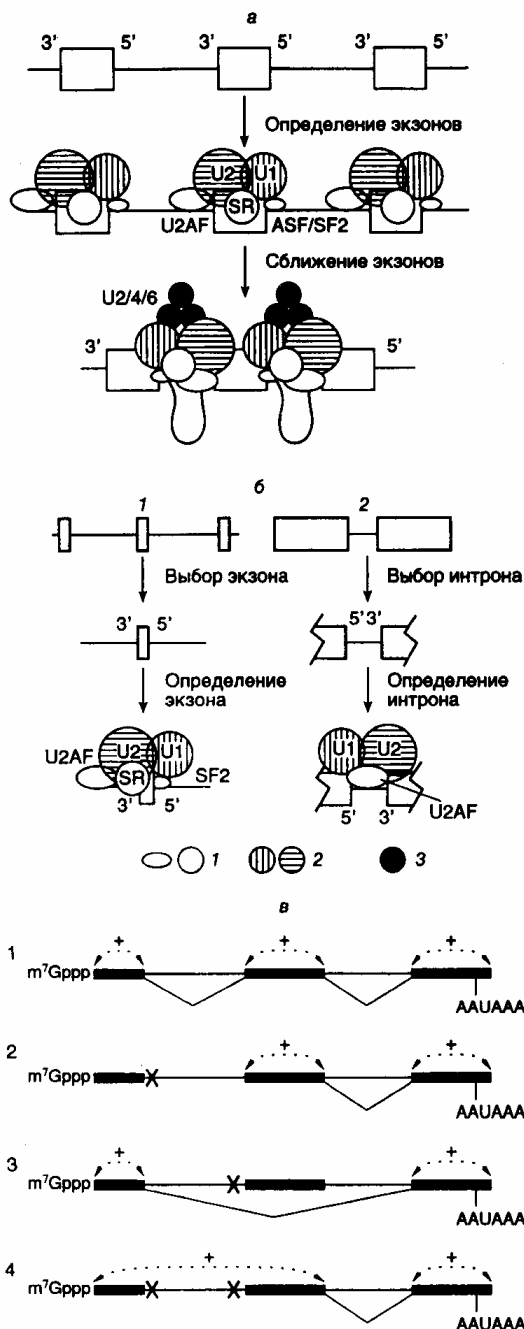


Рис. 1.16. Механизм определения экзонов и интронов при сплайсинге ядерных пре-мРНК

а – определение экзонов; **б** – определение интронов у позвоночных (1) и низших эукариот (2), **в** – роль кэп-связывающего комплекса (CBC) в определении экзонов. Отмечены белки и РНК, участвующие в этом процессе в составе сплайсом: 1 – белки, 2 и 3 – мРНК, X – мутации,

инактивирующие сайты сплайсинга; + – взаимодействия, обеспечивающие распознавание экзонов; AAUAAA – поли(A)-сайт

определения экзонов у животных, взаимодействие между соседними сайтами сплайсинга в генах с короткими экзонами и длинными интронами происходит не через интроны, а через последовательности экзонов. Предполагается, что у пре-мРНК с длинными интронами система сплайсинга, прежде всего, отыскивает пару близкорасположенных сайтов сплайсинга, фланкирующих короткие последовательности экзонов (рис. 1.16,а). После распознавания такой пары с ними взаимодействуют U1- и U2-мяРНП и ассоциированные факторы сплайсинга, включая факторы, распознающие 3'-концевые сайты сплайсинга – U2AF и SC35, а также фактор, узнающий 5'-концевой сайт сплайсинга – ASF/SF2. После завершения процесса определения экзона соседние экзоны входят в контакт друг с другом в результате взаимодействия между факторами, распознающими индивидуальные экзоны. Таким образом, в соответствии с этой моделью процесс сборки активной сплайсома у позвоночных проходит в два этапа, включающие определение экзонов и их сближение между собой. В генах с небольшими интронами, например у низших эукариот, по-видимому, реализуется альтернативный механизм (см. рис. 1.16,б). В этом случае первым этапом сборки сплайсома является *определение интрона*.

В результате мутаций, связанных с нарушением сплайсинга, возникает, по крайней мере, четыре фенотипа: игнорирование экзона, который вырезается вместе с интронами; активация новых (криптических) сайтов сплайсинга; возникновение внутри интронов псевдоэкзонов, последовательности которых не вырезаются из пре-мРНК, и игнорирование интронов. Частоты этих мутаций составляют соответственно 51, 32, 11 и 6%. Все фенотипы, за исключением последнего, могут быть объяснены на основе обсуждаемой модели определения экзонов.

Эта же модель позволяет предсказывать максимальные и минимальные размеры экзонов в генах эукариот. Действительно, анализ длин 1600 внутренних экзонов показал, что только 3,5% из них содержат более 300 нуклеотидов и менее 1% – 400 нуклеотидов. Кроме того, доля экзонов, длина которых меньше 50 нуклеотидов, в этой выборке также незначительна. Известно, что в опытах *in vitro* сплайсинг резко ингибируется, если длина

внутренних экзонов превышает 400 нуклеотидов или становится меньше 50 нуклеотидов. Все эти факты делают модель определения экзонов весьма правдоподобной. Более короткие экзоны, длина которых составляет 6 или 7 п.о., часто обнаруживаемые в генах белков мышц, распознаются системой сплайсинга через энхансерные регуляторные последовательности, расположенные в интронах рядом с экзонами. Механизм распознавания очень длинных экзонов остается неизвестным.

2.2.4. Кэп-связывающий комплекс в роли фактора, сопрягающего основные реакции метаболизма транскриптов РНК-полимеразы II

РНК не может находиться *in vivo* в свободном виде. На протяжении всего внутриклеточного существования – от инициации биосинтеза до полной деградации – РНК пребывает в составе рибонуклеопротеиновых комплексов (РНП). Интенсивные исследования РНП-частиц выявили множество белков, образующих такие комплексы. Поскольку почти все мРНК эукариот претерпевают одни и те же внутриклеточные превращения, основными из которых являются кэпирование, полиаденилирование и сплайсинг, в составе РНП обнаруживают ограниченный набор белков, обеспечивающих протекание этих процессов. Помимо белков, необходимых для специфической трансляции мРНК определенных видов, а следовательно, и ассоциированных только с этими видами мРНК, два белка представлены почти во всех цитоплазматических мРНП животных в больших количествах – уже обсуждавшийся выше поли(А)-связывающий белок РАВР и белок р50. Роль последнего как универсального регулятора трансляции обсуждается в разделе 3.4. Здесь же будут рассмотрены другие, не менее универсальные белки мРНП, которые взаимодействуют с кэп-группами и обеспечивают сопряжение основных реакций посттранскрипционных модификаций мРНК эукариот и экспорта мРНП из ядра в цитоплазму.

Ядерный кэп-связывающий комплекс (СВС). Присоединение кэп-группы к любой строящейся цепи мРНК эукариот, что является ее первой (котранскрипционной) модификацией, оказывает глобальное влияние на все последующие реакции внутриклеточного метаболизма мРНК, включая ее трансляцию. Наличие кэп-группы дает возможность мРНК специфически взаимодействовать с белковым ядерным кэп-связывающим комплексом СВС

(cap-binding complex). Очищенный до гомогенности CBC представляет собой гетеродимер белков CBP80 (молекулярная масса 80 кДа) и CBP20 (20 кДа). N-Концевая часть большой субъединицы CBC содержит сигнал ядерной локализации NLS (nuclear localization signal), обеспечивающий челночные функции CBC во время экспорта мРНК из ядра в цитоплазму (см. ниже). Для взаимодействия CBC с кэп-группой мРНК требуются обе его субъединицы, которые по отдельности не обладают кэп-связывающими свойствами. Обсуждаемые функции CBC сохраняются у всех исследованных видов эукариот – от дрожжей до человека.

Роль CBC в сплайсинге пре-мРНК. Уже более 15 лет известно, что кэп-структура m⁷G необходима для эффективного сплайсинга пре-мРНК в бесклеточных экстрактах клеток животных. Было установлено, что сплайсинг пре-мРНК с одним интроном подавляется в присутствии экзогенных аналогов кэпа или кэп-содержащих РНК. Резкое ингибирование сплайсинга наблюдали в экстрактах, истощенных с помощью специфических антител по большой субъединице CBC – CBP80. Дальнейшее всестороннее изучение этого явления показало, что гетеродимерный CBC требуется для эффективного сплайсинга кэпированных пре-мРНК.

Оказалось, что во время сплайсинга больших пре-мРНК кэп-группа оказывает влияние лишь на вырезание ближайшего к ней интрона. На рис. 1.16в схематически представлены результаты опытов, особенно четко демонстрирующих данное явление. В этих экспериментах использовали искусственные кэпированные пре-мРНК с двумя интронами, которые содержали один или несколько мутантных канонических сайтов сплайсинга в разных комбинациях. Инактивация 5'-концевого сайта сплайсинга первого интрона не оказывала влияния на вырезание второго интрона и блокировала вырезание первого (см. рис. 1.16, в, 2). Мутация в ближайшем к кэп-группе 3'-концевом сайте сплайсинга сопровождалась вырезанием большой последовательности, включающей оба интрона и второй экзон, расположенный между ними (см. рис. 1.16, в, 3). Такая конструкция процессировалась как пре-мРНК, содержащая только один интрон, вырезание которого зависело от наличия комплекса CBC. В том случае, если были инактивированы как 3'-, так и 5'-концевые сайты сплайсинга первого интрона, вырезание второго интрона происходило независимо от CBC (в истощенных по комплексу экстрактах) (не отражено на

рисунке). Лишь одновременное разрушение 5'-концевого сайта сплайсинга и полипиримидиновой последовательности первого экзона сопровождалось появлением СВС-зависимости вырезания второго интрона (см. рис. 1.16, в, 4). (Полипиримидиновая последовательность обсуждалась выше, в начале подраздела, посвященного сплайсингу.) Полученные данные показывают, что именно полипиримидиновая последовательность, а не 5'-концевой сайт сплайсинга первого экзона, обеспечивает кэп-независимое вырезание второго интрона. Принимая во внимание, что СВС необходим для эффективного взаимодействия U1-мяРНП с 5'-концевым сайтом сплайсинга ближайшего к нему интрона, высказывают предположение, что полипиримидиновая последовательность, расположенная перед центральным экзоном, выполняет аналогичные функции при определении этого экзона системой сплайсинга.

Сопряжение сплайсинга с формированием 3'-концевых последовательностей пре-мРНК. Для предшественников мРНК, не содержащих кэп-групп, характерна малая эффективность процессинга их 3'-концевых последовательностей в экстрактах ядер. Кроме того, в таких бесклеточных системах расщепление экспонированной пре-мРНК вблизи поли(А)-сайта подавляется аналогом кэп-группы m^7GpppG и не происходит в экстрактах, истощенных по СВС с помощью антител. На основании такого рода данных делается вывод, что во время процессинга пре-мРНК СВС пространственно сближен с комплексом белков системы, формирующей 3'-концевые последовательности, и стабилизирует этот комплекс через белок-белковые взаимодействия, что необходимо для его эффективного функционирования. Пространственное сближение 5'- и 3'-концевых последовательностей было прямо продемонстрировано для РНП гигантских (35–40 т.о.) транскриптов колец Бальбиани.

В соответствии с современной моделью, поли(А)-сайт пре-мРНК участвует в определении ее последнего интрона вместо обычно используемого для этой цели 5'-концевого сайта сплайсинга. Недавно было установлено, что, в свою очередь, и 3'-концевой сайт сплайсинга последнего интрона стимулирует полиаденилирование пре-мРНК, а следовательно, и процесс удаления последнего интрона сопряжен с полиаденилированием предшественника.

Участие СВС в экспорте транскриптов РНК-полимеразы II из ядер в цитоплазму. Транскрипты РНК-полимеразы II разделяют, по крайней мере, на

два больших класса: мРНК и U-мяРНК. Для РНК обоих классов характерно наличие кэп-групп, но малые ядерные РНК не полиаденилированы. Наличие кэп-группы резко стимулирует экспорт этих РНК из ядер в цитоплазму. С помощью иммуноэлектронной микроскопии удалось осуществить наблюдения *in situ* за динамикой ассоциации, экспорта, а также диссоциации субъединиц СВС и транскриптов колец Бальбиани, которые представляют собой гигантские пuffs на политенных хромосомах слюнных желез комара *Chironomus tentans*. Проведение прямых наблюдений стало возможным благодаря большому размеру транскриптов колец Бальбиани (35–40 т.п.о.) и, как следствие, их премРНП, в состав каждого из которых входит до 200 молекул белков. Было установлено, что белки СВС выходят из ядер в составе мРНП и повторно импортируются в ядра. При этом, в соответствии с современной моделью, происходит следующая последовательность реакций.

Как уже упоминалось, большая субъединица СВС СBP80 содержит специфическую аминокислотную последовательность – сигнал ядерной локализации (NLS). Гетеродимерный цитоплазматический рецептор этого сигнала животных состоит из двух субъединиц – *импортину α* (молекулярная масса 60 кДа) и *импортину β* (90 кДа). В цитоплазме импортин α взаимодействует непосредственно с NLS белков, импортируемых в ядро, и импортин β, который в свою очередь, содержит сайт взаимодействия с ядерным поровым комплексом NPC (nuclear pore complex). Через NPC происходит обмен макромолекулами ядра и цитоплазмы. Белки, содержащие NLS, в комплексе с рецептором переносятся в ядро, где рецептор диссоциирует на субъединицы. Диссоциация рецептора на две субъединицы импортину в нуклеоплазме запускается особым ферментом, GTPазой Ran, связанной с GTP, сродство к которой импортину β выше, чем к импортину α. В результате импортин β удаляется из комплекса. В настоящее время не исключается возможность того, что диссоциировавшая субъединица импортину β остается связанной с цитоплазматической частью NPC. Механизм дальнейшей диссоциации импортину α и NLS не ясен. Полагают, что выход импортину β из комплекса ослабляет связь импортину α и NLS большинства ядерных белков, но не большой субъединицы СВС, с которой импортин α взаимодействует особенно прочно.

В ядре синтезированные РНК через кэп-группу взаимодействуют с СВС, находящимся в комплексе с импортином α , и через NPC перемещаются из ядра в цитоплазму. На цитоплазматической части NPC или в цитоплазме к комплексу СВС–РНК–импортин α присоединяется импортин β через импортин β -связывающий домен импортина α . Это облегчается тем, что в цитоплазме GTPаза Ran находится в комплексе с GDP из-за присутствия так называемого *белка 1, активирующего GTPазу Ran* (Ran GAP1), которая в его присутствии гидролизует связанный с ней GTP. Благодаря особым свойствам NLS кэп-связывающего комплекса СВС, присоединение импортина β в данном случае сопровождается диссоциацией РНК и белкового комплекса СВС-рецептор. В таком виде белковый комплекс переносится в нуклеоплазму, и цикл экспорта синтезированной РНК начинается сначала. В случае U-гяРНП такая цитоплазматическая диссоциация комплекса приводит к гиперметилированию кэп-группы гяРНК, сборке нового комплекса и реимпорту его вместе с гяРНК в ядро. Освобождение кэп-группы мРНК после ее выхода из ядра в цитоплазму делает возможным взаимодействие с ней фактора инициации трансляции eIF-4E с последующим вовлечением мРНК в синтез белка (см. раздел 2.5.1).

Участие СВС в сборке РНП-частиц и формировании правильной пространственной структуры мРНК. Большинство информации о сборке РНП-частиц и механизмах формирования пространственной структуры мРНК в цитоплазме получено с использованием транскриптов колец Бальбиани. Поэтому остается под сомнением возможность распространения выявленных закономерностей на другие транскрипты.

Сборка РНП-частиц осуществляется котранскрипционно (т.е. одновременно с транскрипцией) и начинается с 5'-конца растущей цепи РНК. Поскольку присоединение СВС к кэп-группе происходит одним из первых, полагают, что это инициирует упорядоченное взаимодействие с РНК других белков, входящих в состав гяРНП, с образованием РНП-фибриллы. Наличие кэп-группы дает возможность соответствующим механизмам функционально различать транскрипты РНК-полимераз I, II и III. Действительно, РНК, синтезируемые РНК-полимеразами I и III, не содержат на своих 5'-концах m^7G -групп, включаются в РНП в составе других белков и не подвергаются сплайсингу и полиаденилированию. Для этих транскриптов характерны и другие механизмы транспорта из ядра в цитоплазму. Полагают, что набор белков,

ассоциированных с РНК разных классов, определяется, прежде всего, особенностями их первичной структуры. В то же время белки РНП обеспечивают РНК пространственную структуру, оптимальную для ее дальнейшего внутриклеточного метаболизма и функционирования.

2.3. Функциональная компартментализация ядра

При рассмотрении механизмов реализации генетической информации на уровне транскрипции и посттранскрипционных модификаций РНК чаще всего не принимается во внимание пространственная внутриклеточная организация макромолекулярных комплексов, обеспечивающих осуществление этих механизмов. Имеются, по крайней мере, две причины, по которым такого рода вопросам в современных исследованиях уделяется недостаточное внимание. Во-первых, большинство молекулярно-генетических исследований механизмов транскрипции проводится в бесклеточных системах, где вполне удовлетворительно выявляются отдельные компоненты функциональных макромолекулярных комплексов и их биохимические функции. Во-вторых, такие эксперименты сопряжены со значительными методическими сложностями и в настоящее время, как правило, ограничиваются различными приложениями микроскопии в сочетании с молекулярными зондами. Однако переход с фундаментального молекулярного уровня исследований в бесклеточных системах ("метаболического котла") на более высокий, учитывающий пространственную организацию биохимических процессов, неизбежен, так как именно на нем реализуются жизненно важные принципы интеграции и регуляции метаболизма живого организма во всех его проявлениях.

Концепция функциональной компартментализации эукариотической клетки в настоящее время общепринята и основывается на признании существования многочисленных внутриклеточных органелл, одной из которых является ядро. Наличие дискретных микрокомпарментов в отдельных органеллах менее очевидно, однако реально подтверждается в эксперименте. Результаты, полученные при использовании нескольких независимых подходов, указывают на компартментализацию процессов репликации, транскрипции и процессинга РНК в эукариотических ядрах. Так, синтез ДНК и РНК тонко организован во времени и пространстве. Интерфазные хромосомы занимают в ядрах вполне определенные микрокомпарменты (хромосомные

зоны или территории) и не перемешаны друг с другом. При этом их пространственная структура высокоупорядочена.

2.3.1. Интерфазные хромосомы в ядре

В разделе 1.3 уже кратко обсуждался петельно-доменный уровень структурной организации хромосом эукариот, который отражает разделение интерфазных хромосом на дискретные домены по функциональному признаку. Наибольшее число экспериментальных данных в пользу наличия таких функциональных доменов было получено в исследованиях хромосом дрозофилы в связи с мозаичным эффектом положения, в которых использовались методы генетического анализа в сочетании с цитологическими наблюдениями политенных хромосом. Феномен эффекта положения и молекулярные механизмы, лежащие в его основе, будут более подробно рассмотрены в разделе 3.2.4. Говоря коротко, эффект положения заключается в том, что перенос активно транскрибируемых генов к неактивным гетерохроматизированным участкам хромосом часто сопровождается значительным подавлением транскрипционной активности генов в их новом положении на хромосоме. Возможен поиск генов, ингибирующих или усиливающих эффект положения генов-репортеров, уровень экспрессии которых легко устанавливается фенотипически, например по изменению интенсивности окраски глаз. Многие из идентифицированных таким образом генов кодируют структурные компоненты хроматина или ферменты, ковалентно модифицирующие эти компоненты.

Эффект положения рассматривается в настоящее время в качестве универсального явления, характерного для всех эукариотических хромосом. Он наглядно демонстрирует наличие в интерфазных хромосомах высокоупорядоченной доменной структуры, тесно связанной с транскрипционной активностью соответствующих участков генома. Распределение участков хроматина с разным уровнем конденсированности в интерфазных хромосомах упорядочено и, по-видимому, является видовым признаком организма. Имеются указания на то, что структурные компоненты конденсированного хроматина оказывают влияние на пространственное расположение соответствующих частей хромосом в ядре. Следовательно, такая их микрокомпартиментализация может играть важную роль в регуляции

экспрессии генов.

Уже в ранних цитологических экспериментах (С. Рабл, 1885 г.) было показано, что теломерные участки хромосом клеток слюнных желез саламандр располагаются вблизи ядерной оболочки. В настоящее время установлена локализация на периферии ядер теломер политенных хромосом дрозофилы, а также теломерных участков хромосом *Schizosaccharomyces pombe* в фазе G₂ клеточного цикла. С помощью конфокальной иммунофлуоресцентной микроскопии продемонстрировано, что перинуклеарную локализацию теломер дрожжей *S. cerevisiae* обеспечивают, по крайней мере, два белка – Sir3 и Sir4 (silent information regulators), которые также требуются для наследуемой инактивации генов в специфических доменах хромосом, расположенных в локусе, определяющем тип спаривания у дрожжей, и вблизи теломерных последовательностей. В такого рода экспериментах была установлена связь между ядерной локализацией конкретных участков хромосом и их транскрипционной активностью.

В отличие от теломерных участков хромосом, активно экспрессирующиеся гены локализуются преимущественно во внутренних частях интерфазных ядер. При этом, по мнению Д.Б. Лоуренса и соавторов (1993 г.), отдельные гены внутри ядер располагаются упорядоченно. В частности, было установлено, что активно транскрибируемые гены вируса Эпштейна–Барр и онкогена *neu* находятся в разных местах внутренних 50% ядерного объема, а ген дистрофина – вблизи ядерной оболочки. В этой серии экспериментов три неактивных гена, кодирующих альбумин, тяжелую цепь сердечного миозина и нейротензин, локализовали в составе конститутивного гетерохроматина на периферии ядер или вблизи ядрышка. Используя микрооблучение ультрафиолетовым светом и гибридизацию с зондами, показали, что отдельные хромосомы занимают внутри ядра дискретные, хотя и обширные, территории. Компоненты аппарата сплайсинга обнаруживают в ядрах на периферии территорий, занимаемых индивидуальными хромосомами, так же как и треки синтезируемой РНК (см. ниже). В соответствии с моделью Т. Кремера (1993 г.), внутриядерное пространство между территориями, занимаемыми индивидуальными хромосомами, представляет собой единый компартмент, в котором происходят транскрипция, сплайсинг, созревание транскриптов и их транспорт. Этот компартмент тесно ассоциирован с активно

транскрибируемыми генами, которые располагаются на периферии хромосомных территорий в составе выступающих петель хроматина. Несмотря на большую функциональную важность обсуждаемого вопроса, истинная природа интерфейса между активными генами отдельных хромосомных территорий и межхроматиновым компартментом остается неясной из-за слабого понимания пространственной структуры ДНК на высших уровнях ее упаковки в хромосомах.

2.3.2. Ядрышко

Структурно-функциональная организация *ядрышка* (nucleolus) еще более наглядно иллюстрирует концепцию функциональной компартментализации ядра эукариотических клеток. В этой части ядра происходят транскрипция рибосомных генов, процессинг предшественников рРНК и сборка прерибосомных частиц из рибосомных белков и рРНК. Механизмы формирования ядрышка, приводящие к образованию внутри ядра этой дискретной и легко обнаруживаемой структуры, в настоящее время не ясны. В соответствии с одной из простых гипотез, ядрышко рассматривают как нуклеопротеиновый комплекс, спонтанно появляющийся в результате объединения регуляторных белково–нуклеиновых комплексов, возникающих на повторяющихся последовательностях рДНК во время их транскрипции. Действительно, гены рРНК человека организованы в виде 250 tandemно повторяющихся последовательностей длиной в 44 т.п.о. каждая, которые вместе с ассоциированными с ними белками формируют сердцевину ядрышка. Оно заполняется другими компонентами во время процессинга рРНК и сборки рибосомных субчастиц.

Морфологически в ядрышке различают три основные зоны: фибриллярный центр (1), окруженный плотной фибриллярной (2) и гранулярной (3) областями. С помощью специфических антител и гибридизационных зондов было установлено, что в фибриллярном центре ядрышка локализованы гены рРНК, РНК-полимераза I, транскрипционный фактор UBF и топоизомераза I. Полагают, что фибриллярный центр ядрышка является местом сборки регуляторных нуклеопротеиновых комплексов, необходимых для транскрипции генов рРНК. Плотный фибриллярный компонент, окружающий центр ядрышка, представлен растущими цепями предшественников рРНК и ассоциированными

с ними белками, участвующими в процессинге. В гранулярной области ядрышка обнаруживают зрелые 28S и 18S рРНК, частично процессированные РНК, а также продукты сборки рибосомных субчастиц. Интермедиаты сборки рибосом представлены частицами диаметром 15–20 нм. Перенос прерибосомных субчастиц к цитоплазме, по-видимому, обеспечивают специфические белки, которые направленно перемещаются от ядрышка к оболочке ядра. Благодаря четко прослеживаемой иерархии в структурно-функциональной организации ядрышка в виде отдельных морфологически различимых компартментов его часто используют в качестве модели функциональной компартментализации синтеза мРНК, ее процессинга и экспорта в цитоплазму.

При этом следует иметь в виду, что наблюдаемая "высокоупорядоченная" пространственная структура ядрышка может быть просто следствием функционирования большого числа генов рРНК, организованных в тандемные повторы, что сопровождается накоплением транскриптов РНК-полимеразы I и продуктов их процессинга в окрестностях активно работающих генов. Структура ядрышка является динамической, а его пространственное расположение и структурные особенности зависят от внутриядерной локализации и уровня активности соответствующих генов рРНК.

Долгое время ядрышко рассматривалось только в обсуждаемом выше аспекте, т.е. как внутриядерный микрокомпартмент биогенеза рибосом, в котором происходят транскрипция рДНК и сборка рибосомных субчастиц из составляющих компонентов. Однако в последнее время начинают появляться данные, указывающие на участие ядрышка в регуляции клеточного цикла.

Даже геном дрожжей содержит ~200 тандемно повторяющихся генов рРНК. При этом не все гены одинаковы в функциональном отношении: транскрибируется лишь половина последовательностей рДНК, а в их воспроизводстве задействовано лишь ~20% имеющихся областей начала репликации. Перенос генов в область рДНК часто сопровождается их репрессией, что, как полагают, является следствием функционирования механизма подавления гомологичной рекомбинации в участках генома, содержащих тандемные повторы. Мутационное нарушение этого механизма сопровождается образованием сотен внехромосомных кольцевых рДНК, которые неравномерно распределяются между дочерними клетками во время митоза. Накопление материнскими клетками внехромосомных рДНК приводит к

уменьшению способности клеток делиться. Этот феномен был назван "старением клеток" (cellular aging). Кроме того, складывается впечатление, что ядрышко может регулировать входение клеток в мейоз, а также активность фосфатазы Cdc 14, контролирующей прохождение телофазы митоза. Получены данные о том, что повторяющиеся последовательности рДНК ядрышка служат местом сборки большого регуляторного белкового комплекса RENT (regulator of nucleolar silencing and telophase exit), в состав которого входит вышеупомянутая фосфатаза и, как минимум, три других белка, которые и обеспечивают регуляторные функции ядрышка.

2.3.3. Пространственная организация синтеза мРНК

Внутриядерный синтез мРНК и доставка зрелых транскриптов к месту их трансляции требуют участия множества тонко сбалансированных во времени, пространственно организованных молекулярных механизмов. Выше уже были рассмотрены основные молекулярные процессы, обеспечивающие сборку инициационных и элонгирующих нуклеопротеиновых комплексов, а также механизмы котранскрипционных и посттранскрипционных модификаций РНК, включая кэпирование, сплайсинг, редактирование их первичной структуры и полиаденилирование. Теперь кратко суммируем известные факты о внутриядерной компартментализации этих процессов.

Аппарат транскрипции, участвующий в синтезе пре-мРНК, ассоциирован с *перихроматиновыми фибриллами*, обнаруживаемыми на границах доменов конденсированного хроматина. Эти фибриллы представляют собой ядерные рибонуклеопротеиновые комплексы диаметром 3–20 нм. Они включают в себя растущие цепи пре-мРНК, и плотность фибрилл коррелирует с транскрипционной активностью соответствующих участков хроматина. С помощью иммунохимических методов здесь же обнаружены компоненты аппарата сплайсинга, который удаляет интроны из предшественников мРНК одновременно с элонгацией транскриптов.

В опытах по внутриядерной локализации мест синтеза специфических транскриптов с использованием импульсной радиоактивной или флуоресцентной меток такие РНК были обнаружены в виде "треков" или более компактных "точек" в одном или двух дискретных участках ядра, что соответствует копиям соответствующих генов на гомологичных хромосомах.

При этом с использованием одновременной гибридизации ДНК и РНК было показано, что транскрибируемые гены расположены прямо в треках или точках на одном из концов трека. Более того, зонды, специфичные в отношении последовательностей интронов, метят треки только вблизи гена, указывая на то, что сплайсинг происходит вдоль этого следа РНК.

Треки РНК тесно ассоциированы с дискретными внутриядерными структурами, называемыми *межхроматиновыми гранулами*, или *спеклами*. Спеклы обогащены компонентами аппарата сплайсинга, а также содержат интронсодержащие пре-мРНК и полиаденилированные молекулы. В соответствии с этим спеклы могут представлять собой места процессинга пре-мРНК и аккумуляции зрелых мРНК внутри ядер. В ряде случаев выявляется неслучайная ассоциация активно транскрибируемых генов со спеклами, которые могут маркировать внутриядерные области транскрипции. Противоречивость этой интерпретации заключается в том, что в ядре обнаруживаются одновременно всего 20–50 спеклов, тогда как транскрибирующихся генов значительно больше. Следовательно, не каждый транскрибируемый ген ассоциирован с такими структурами. Большинство экспериментальных доказательств ассоциации мест транскрипции со спеклами получено для особо активных генов, например гена коллагена, транскрипты которого в фибробластах составляют до 4% суммарной РНК. Возможно, спеклы представляют собой микрокомпартменты, в которых происходит процессинг РНК наиболее активно транскрибируемых генов.

После завершения синтеза и объединения с компонентами аппарата сплайсинга, формирующими сплайсомы, пре-мРНК переносится к ядерной оболочке и выходит в цитоплазму. Выход РНК из ядра в цитоплазму осуществляется через поры в ядерной оболочке, входящие в состав ядерного порового комплекса (NPC). NPC является постоянно существующим микрокомпартментом, однозначно идентифицируемым с помощью микроскопических, генетических и биохимических методов.

Во время внутриядерного транспорта молекулы РНК объединяются с другими белками, участвующими в их процессинге, образуя *гетерогенные ядерные рибонуклеопротеиновые комплексы* (гяРНП), в которых пространственная структура пре-мРНК оптимизирована для ее созревания. Ядерный матрикс, получаемый в результате удаления из ядер большей части

хроматина, все еще содержит пре-мРНК, гяРНП и некоторые компоненты аппарата сплайсинга. Это может указывать на внутриядерную избыточность гяРНП и их роль в формировании пространственной структуры ядер. В общем, для гяРНП характерно диффузное распределение в нуклеоплазме, однако часть из них может концентрироваться в окрестностях спеклов и даже следовать за РНК из ядра в цитоплазму и вновь возвращаться в ядра вместе с транскриптами.

Процесс перемещения специфических пре-мРНК от генов в цитоплазму клеток был подробно исследован в случае экспорта частиц пре-мРНК колец Бальбиани комара *Chironomus tentans*, экспрессирующие гены которых находятся в составе гигантских пуффов политенных хромосом слюнных желез. Крупные транскрипты этих пуффов во время элонгации упаковываются в гяРНП в виде тонких фибрилл, которые по мере удлинения пре-мРНК становятся толще и изгибаются с образованием кольцеобразных структур. Зрелые гранулы пре-мРНК, которые, как полагают, заключают в себе РНК, претерпевшую сплайсинг, движутся в нуклеоплазме к ядерной оболочке, где задерживаются вблизи ядерных пор. В это время они приобретают форму палочек, которые проходят через ядерные поры, начиная с 5'-конца заключенной в них РНК. Как только пре-мРНК появляются на поверхности цитоплазматической части ядерной мембраны, они объединяются с рибосомами. Следует подчеркнуть, что на протяжении всей этой цепи событий РНК находится в составе пространственно упорядоченных РНП-частиц. Накапливаются данные в пользу того, что перемещение РНК от гена к ядерной мембране не является следствием простой диффузии. Такому простому объяснению, в частности противоречат факты тесной ассоциации транскриптов, гяРНП и компонентов аппарата сплайсинга с ядерным матриксом. В ряде случаев находит подтверждение гипотеза ядерной фиксации генов (*gene gating model*), предложенная Г. Блобелом (1985 г.), в соответствии с которой конкретные гены функционально связаны с определенными участками (и порами) ядерной мембраны, что направляет их транскрипты для экспорта в цитоплазму к этим конкретным участкам. Однако такое правило подтверждается не всегда. В частности, РНК коллагена обнаруживают распределенной вдоль всей ядерной мембраны, что указывает на ее выход в цитоплазму через многие ядерные поры.

Для рассмотренных выше структур, образование которых сопровождается синтезом, процессингом и экспортом РНК из ядра в цитоплазму, характерен динамизм – отдельные их компоненты могут перемещаться между микрокомпартаментами. Синтез, процессинг и транспорт РНК в ядре происходят в составе дискретных компартментов нуклеоплазмы, что позволяет концентрировать регуляторные, структурные и ферментативные компоненты транскрипции и сплайсинга в местах активно экспрессирующихся генов. Действительно, все этапы сплайсинга можно воспроизвести в разбавленных растворах *in vitro* при концентрации белка ~1 мкг/мл. Однако скорость этих реакций в таких системах значительно ниже наблюдаемой *in vivo*, где внутриядерная концентрация РНП превышает 50 мг/мл. Кроме того, пространственно упорядоченная организация ранних этапов экспрессии генов создает необходимые условия и дополнительные уникальные возможности для ее регуляции, что было бы невозможно в случае свободной диффузии компонентов этой системы.

2.3.4. Ядерные тельца и домены

Исследования структурно-функциональных отношений в ядре в связи с компартментализацией транскрипции, процессинга РНК и репликации продемонстрировали наличие особых функций у многих морфологически различных внутриядерных микроструктур. В этом отношении не явились исключением и *ядерные тельца*, вначале описанные чисто морфологически.

Свернутые тельца (coiled bodies). Эти внутриядерные ультраструктуры, обнаруживаемые в виде клубка переплетенных нитей в ядрах клеток млекопитающих, часто ассоциированы с периферией ядрышка. В составе этой ультраструктуры обнаружены белки и РНК ядрышка, участвующие в модификации и процессинге рРНК, включая *фибрилларин* и малые ядрышковые РНК U3. В этих структурах также обнаруживают малые ядерные РНП, в частности U7-мРНК, и некоторые специализированные белки (например *койлин* р80). Все это указывает на возможное участие свернутых телец в процессинге РНК. Для данных ультраструктур характерен динамизм: они исчезают в митозе и вновь возникают в фазе G₁ клеточного цикла. Их количество резко возрастает в условиях стимуляции пролиферации клеток.

Аналогами свернутых телец являются *снурпосомы* С ядер ооцитов

амфибий. Эти тельца ассоциированы с локусом гистоновых генов хромосом типа ламповых щеток и обогащены U7-мРНК, которая вовлечена в модификацию 3'-концов гистоновых пре-мРНК. Высказывается предположение, что и снурпосомы, и свернутые тельца могут участвовать в посттранскрипционных модификациях гистоновых и других пре-мРНК.

Ядерные тельца PML. Острая промиелоцитарная лейкемия (PML) часто ассоциирована с транслокацией t(15;17) (перенос участка хромосомы 15 на хромосому 17), которая приводит к слиянию в одной рамке считывания гена PML с геном α -рецептора ретиноевой кислоты. У нормальных индивидуумов белок PML, обладающий Zn^{2+} -связывающим доменом типа "пальцы RING", локализован в отдельном ядерном микрокомпартементе в виде плотного фибриллярного кольца, окружающего сердцевину. Этот белок обнаруживают также вместе с U1-мРНК и койлином p80 в особых зонах, окружающих межхроматиновые гранулы, или спеклы. Ядерные тельца PML претерпевают морфологические изменения на протяжении клеточного цикла. Эти тельца разрушаются во время вирусной инфекции, а для репликации аденовирусной ДНК их разрушение является необходимым этапом, что подчеркивает возможное участие телец в обеспечении антивирусной активности клеток. Инкубация клеток с интерфероном индуцирует синтез PML-белка и подавляет размножение вирусов.

У пациентов с промиелоцитарной лейкемией, содержащих вышеупомянутый гибридный ген, имеет место разрушение ядерных телец PML. Инкубация лейкозных клеток с ретиноевой кислотой приводит к восстановлению этих морфологических структур, что коррелирует с наступлением ремиссии у больных острой промиелоцитарной лейкемией. Предполагается, что тельца PML дикого типа участвуют в подавлении неконтролируемого роста трансформированных клеток и понижают уровень их злокачественности по непонятному пока механизму.

Ядерные домены WT1. Ген *WT1* человека является геном-супрессором опухолей, мутационные нарушения которого наблюдают при злокачественных новообразованиях Вилмса, возникающих в почках в детском возрасте. Четыре белка, кодируемые этим геном, возникают в результате альтернативного сплайсинга, и, по крайней мере, один из них является фактором транскрипции. Для полноразмерного белка характерно наличие С-концевого мотива типа

"цинковых пальцев" (см. раздел 2.2.2) и N-концевого домена, обогащенного Pro/Gln. Белок WT1 обладает способностью связываться со специфической GC-богатой последовательностью нуклеотидов и подавлять транскрипцию генов, промоторы которых содержат эту последовательность. Для него характерна дискретная внутриядерная локализация в составе морфологически различаемых телец. Ассоциация изоформ WT1 с аппаратом сплайсинга предполагает, помимо прямого влияния на транскрипцию, их участие в посттранскрипционной регуляции экспрессии генов.

2.3.5. Компартментализованное ядро

Два основных структурных образования характерны для ядер всех эукариот. Это, во-первых, оболочка ядра с ядерными порами, связанная с *ядерной ламиной* (электронно-плотный слой, прилегающий к ядерной оболочке со стороны нуклеоплазмы), и, во-вторых, хромосомы. Именно транскрипция и репликация ДНК хромосом являются теми процессами, для обеспечения которых, прежде всего, существует ядро в эукариотической клетке. Как уже было показано выше, зоны транскрипционной активности хромосом компартментализованы внутри ядер. То же характерно и для репликации ДНК эукариот (см. главу 4). Неразрешенным остается вопрос о природе внутриядерных механизмов, которые могли бы лежать в основе такой функциональной компартментализации. В этом случае фактором, осложняющим проведение исследований, является динамизм структуры хроматина, которая претерпевает упорядоченные перестройки во время репликации и транскрипции. То же относится и к пространственной структуре хромосом, изменяющейся на протяжении клеточного цикла.

Молекулярный аппарат, обеспечивающий внутриядерную экспрессию генов, их репликацию и репарацию, настолько велик, что объединение его компонентов во время функционирования может приводить к формированию надмолекулярных комплексов, различимых морфологически. Это особенно наглядно демонстрирует структура ядрышка, в котором можно легко выделить отдельные функциональные домены, содержащие: 1) регуляторные нуклеопротеиновые комплексы; 2) активно функционирующий аппарат транскрипции; 3) рибонуклеопротеиновые комплексы процессируемых транскриптов. Те же домены, хотя и менее масштабные, обнаруживаются при

функционировании РНК-полимеразы II. Многие черты молекулярной организации гигантских молекулярных комплексов, обеспечивающих транскрипцию (транскрипсосомы) и сплайсинг (сплайсомы) у эукариот, в настоящее время уже выяснены и были рассмотрены выше. Интенсивно исследуются молекулярная структура и пространственная организация аппарата репликации (реплисомы), которые будут подробно обсуждаться в главе 4. Однако ясное понимание внутриядерных связей между этими ключевыми комплексами в настоящее время отсутствует.

Следует особо подчеркнуть роль индивидуальных эукариотических хромосом в формировании пространственной структуры ядра. Не исключено, что необходимость организации генетического материала в виде индивидуальных хромосом, число, размер и трехмерная структура которых являются фундаментальными таксономическими признаками, продиктована потребностью соматических клеток в упорядоченном распределении матричной ДНК по внутриядерным территориям с целью координации экспрессии генов и стабилизации заключенной в них генетической информации. Выше уже отмечалась преимущественная локализация теломерных последовательностей нуклеотидов вблизи ядерной оболочки, а также расположение последовательностей наиболее активно транскрибируемых генов во внутренних частях ядер. В соответствии с одной из точек зрения (Д. Строболис, А.П. Волффе, 1996 г.), такая пространственная организация генома может облегчать экспорт синтезирующейся мРНК из ядра в цитоплазму. Однако, на мой взгляд, могут быть и более глубокие причины такой пространственной организации генома эукариот. Гетерогенность хроматина, проявляющаяся в различных уровнях его конденсации в отдельных генетических локусах, а также неслучайное внутриядерное распределение последовательностей генома создают многоуровневую защиту генетической информации от химических мутагенов и могут контролировать скорости изменения отдельных генетических локусов в филогенезе многоклеточных организмов (подробнее см. раздел 5.3). Ядро эукариот обеспечивает прохождение первых этапов реализации генетической информации: избирательную транскрипцию генов, а также посттранскрипционные модификации и процессинг предшественников РНК, которые позволяют им вступать в трансляцию, т.е. реализовывать следующий важнейший этап передачи генетической информации от генов к белкам.

2.4. Биосинтез белка рибосомами бактерий

В процесс биосинтеза белка рибосомами, называемого *трансляцией*, вовлечено множество макромолекул и макромолекулярных комплексов. На этом этапе реализации генетической информации происходит считывание генетической информации, заключенной в мРНК, рибосомами и ее передача полипептидным цепям белков, т.е. биосинтез полипептидных цепей, последовательность аминокислот в которых, как правило, однозначно определена последовательностью нуклеотидов в транслируемых мРНК в соответствии с генетическим кодом. Свободные аминокислоты не узнаются рибосомами. Чтобы это произошло, аминокислоты должны поступать в рибосомы в виде конъюгатов с тРНК (аминоацилированных тРНК), последовательности нуклеотидов которых распознаются аппаратом трансляции. В каждой молекуле тРНК имеется участок из трех нуклеотидов, комплементарный кодону мРНК. Именно эта последовательность, называемая *антикодоном*, в основном определяет положение той или иной аминокислоты в полипептидной цепи. В ходе каждого индивидуального акта трансляции рибосома распознает кодон мРНК и в соответствии с ним выбирает аминоацилированную тРНК, антикодон которой соответствует транслируемому кодону. После этого происходит соединение посредством пептидной связи очередной аминокислоты с С-концевой аминокислотой растущей цепи полипептида.

Таким образом, во время трансляции рибосома после связывания мРНК начинает последовательно, кодон за кодоном, перемещаться вдоль матрицы, выбирая из окружающей среды молекулы аминоацилированных тРНК. При этом каждый индивидуальный акт трансляции завершается присоединением выбранной молекулы аминокислоты к С-концевой аминокислоте синтезируемой цепи белка посредством пептидной связи. Ниже более подробно будут рассмотрены основные этапы биосинтеза белка и компоненты белоксинтезирующей системы бактерий.

2.4.1. Рибосомы

Рибосомы представляют собой крупный рибонуклеопротеидный комплекс с молекулярной массой ~ 2,5 мДа, состоящий из рибосомных белков,

молекул рРНК и ассоциированных с ними факторов трансляции. Рибосомы прокариотических и эукариотических организмов различаются по размерам. У эукариот они представлены 80S частицами, тогда как коэффициент седиментации рибосом прокариот составляет 70S. Рибосомы всех известных организмов построены из двух неравных субчастиц: прокариотические – 30S и 50S, а эукариотические – 40S и 60S. 70S рибосомы эубактерий в своем составе содержат 55–60 рибосомных белков, для 80S рибосом эукариот это число составляет 75–85. В обоих случаях рибосомные белки в составе рибосом ассоциированы с молекулами рРНК, образуя пространственно организованные рибонуклеопротеиновые тяжи.

Рибосомные белки *E. coli*. В настоящее время более 50 рибосомных белков выделено в высокоочищенном состоянии. Молекулярная масса самого маленького белка составляет 5 кДа, а самого большого – 61 кДа, тогда как для большинства рибосомных белков эти значения лежат в пределах 10–20 кДа. Определены аминокислотные последовательности полипептидных цепей всех рибосомных белков *E. coli*. Малая рибосомная субчастица содержит 21 белок с суммарной молекулярной массой 350 кДа.

Белки в составе 30S субчастицы ассоциированы с 16S РНК, длина которой составляет 1542 нуклеотида (нт). Суммарные молекулярные массы малой и большой субчастиц рибосом достигают соответственно 850 и ~ 1450 кДа. Третья часть массы большой субчастицы приходится на 34 рибосомных белка, а две третьих – на 23S (2904 нт) и 5S рРНК (120 нт). Продолжают накапливаться биохимические данные, указывающие на центральную, возможно ключевую, роль рРНК в обеспечении этапов трансляции. Обнаружены специфические внутримолекулярные и межмолекулярные взаимодействия между различными функциональными участками рРНК. На прямое участие 23S рРНК в трансляции указывает наличие специфических комплементарных взаимодействий между ней и ССА-концами тРНК, акцептирующими аминокислотные остатки. В уточненных пространственных моделях 30S и 50S субчастиц, рассмотренных ниже, сегменты рРНК со специфическими структурными особенностями располагаются в функционально значимых участках рибосом.

Методы исследования пространственной структуры рибосом. Вся информация о пространственной структуре рибосом получена с

использованием современных методов, в которых реализуются два направления исследований: структурно-биологические исследования низкого и высокого разрешения, а также биохимические – высокого разрешения. Получаемые результаты, по мере совершенствования методов, все более сближаются друг с другом.

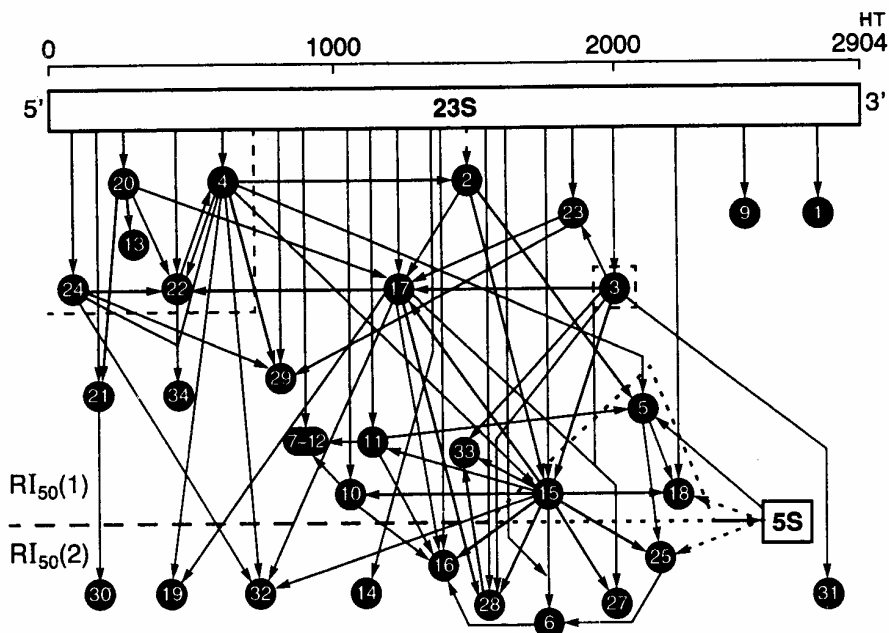


Рис. I.17. Карта сборки 50S субчастицы рибосом *E. coli* из индивидуальных рибосомных белков и рРНК в процессе реконструкции *in vitro*

Стрелками обозначены межмолекулярные взаимодействия, удерживающие индивидуальные белки (цифры в кружках) в составе субчастицы, и взаимозависимость сборки. Толщина стрелок отражает прочность связей

Одними из первых следует упомянуть *кристаллографию* и *ЯМР-спектроскопию*. Имеются значительные достижения в изучении пространственной структуры небольших субдоменов рРНК с помощью ЯМР-спектроскопии, а также в исследовании структуры индивидуальных рибосомных белков с использованием методов ЯМР и кристаллографии. В последнее время закристаллизованы отдельные рибосомные субчастицы и 70S рибосомы галофильных и термофильных бактерий. Получены картины дифракции для 50S субчастиц галофилов с разрешением 3 Å, а также рибосом и субчастиц *T. thermophilus* с разрешением 7–20 Å. Однако эти картины дифракции, по

крайней мере, в 10 раз сложнее тех, с которыми приходилось иметь дело раньше при расшифровке структур несимметричных макромолекул, что сильно затрудняет их интерпретацию. В результате наиболее распространена в настоящее время альтернативная стратегия определения структуры рибосом "по частям": через пространственные структуры индивидуальных рибосомных белков, субдоменов рРНК и их небольших комплексов. Выявленное в ходе этих исследований большое разнообразие пространственных структур рибосомных белков указывает на существование не менее разнообразных механизмов распознавания белками специфических участков РНК, количество которых должно значительно превышать ограниченное число известных механизмов ДНК–белкового узнавания.

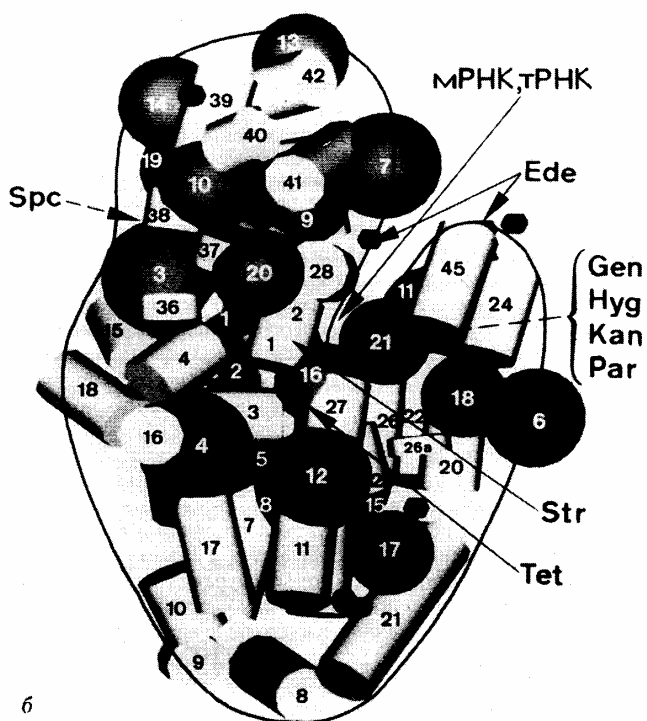
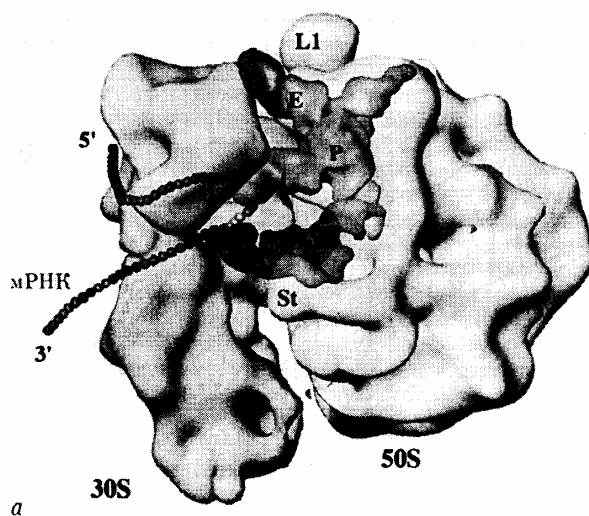


Рис. 1.18. Полупрозрачная модель пространственной структуры рибосомы *E. coli* (а) и расположение рибосомных белков, а также рРНК в ее 30S-субчастице (б)

Указано положение мРНК, тРНК в А-, Р- и Е-участках рибосомы, белка L1 и "выступа" (stalk) 50S субчастицы. Черными шарами обозначены рибосомные белки, светлыми цилиндрами – спиральные участки 16S

рРНК, стрелками – места связывания мРНК, тРНК и антибиотиков, цифрами – рибосомные белки и участки рРНК согласно общепринятой номенклатуре

Другим плодотворным направлением исследования пространственной структуры рибосом при низком разрешении является *электронная микроскопия* (ЭМ). При этом образцы с рибосомами быстро замораживают в жидком этане в тонком слое буфера и исследуют с помощью ЭМ при температуре жидкого азота и низких дозах облучения для сохранения чувствительных к радиации структур в интактном состоянии. На получаемых в результате микрофотографиях могут одновременно содержаться сотни и тысячи поразному ориентированных индивидуальных рибосом, изображения которых далее подвергаются компьютерному анализу с последующей реконструкцией трехмерной структуры индивидуальной рибосомной частицы. Реконструкцию трехмерных структур работающих рибосом получают в результате анализа микрофотографий ультратонких срезов отдельных бактериальных клеток, активно синтезировавших белок или находившихся в состоянии блока трансляции. Одной из разновидностей ЭМ, активно используемой для исследования пространственной структуры рибосом, является *иммуноэлектронная микроскопия*. Первичные и пространственные структуры рибосомных белков, формирующие их эпитопы, значительно различаются, поэтому такие белки редко дают перекрестные иммунологические реакции и их можно четко идентифицировать с помощью специфических антител. При анализе комплексов антител с рибосомами с помощью электронной микроскопии можно видеть, что многие белки локализованы на поверхности рибосомных субчастиц. Оказалось, что пространственное расположение большинства рибосомных белков весьма консервативно. В частности, у грамположительных и грамотрицательных бактерий гомологичные белки занимают одни и те же места на поверхности рибосомных субчастиц.

Важная информация о расположении индивидуальных рибосомных белков в составе рибосом получена и с помощью *метода поперечных сшивок*. Используя бифункциональные реагенты, например диэпоксидбутан или 2-иминотиолан, осуществляют ковалентное соединение рибосомных белков, расположенных по соседству на расстоянии 5–10 Å друг от друга. Поскольку в настоящее время аминокислотные последовательности всех рибосомных

белков известны, этим методом можно однозначно определять, какие аминокислотные остатки в соседних белках участвуют в образовании поперечных сшивок.

Целостность рибосом в водных растворах в значительной степени зависит от температуры и ионных условий, особенно от концентрации двухвалентных ионов (Mg^{2+} и Ca^{2+}). Понижение концентрации ионов Mg^{2+} приводит вначале к диссоциации рибосом на большую и малую субчастицы, а затем к последовательному упорядоченному освобождению рибосомных белков из субчастиц вплоть до их полного распада с образованием пула отдельных белков и рРНК. Процесс разборки рибосом обратим, и при восстановлении ионных и температурных условий в реакционной смеси возможно реконструирование рибосомных субчастиц из отдельных компонентов с образованием полноценных функционально активных рибосом. На рис. 1.17 представлена карта сборки большой субчастицы рибосом *E. coli* из отдельных компонентов, которая отражает последовательность присоединения рибосомных белков к рРНК, а также два основных этапа сборки. Для перехода ко второму этапу необходимо дальнейшее изменение ионных условий и температуры реакционной среды. Процесс сборки субчастиц рибосом является кооперативным, т.е. присоединение одних рибосомных белков стимулирует включение других. При этом белки, включающиеся в состав рибосомных субчастиц друг за другом, в зрелых субчастицах оказываются расположенными рядом. *Реконструирование рибосом из отдельных компонентов in vitro* внесло большой вклад в понимание пространственной организации рибосомных субчастиц и функциональной значимости отдельных рибосомных белков.

Метод рассеяния нейтронов на протонах, входящих в состав белков, также способствовал пониманию пространственного расположения белков в рибосомах. Отклонения нейтронов после контактов с протонами белков можно легко отличить от отклонений, которые являются результатом взаимодействия нейтронов с другими атомами, в частности тяжелыми изотопами водорода (2H или 3H). Если в состав дейтерированных рибосомных субчастиц ввести два рибосомных белка, содержащих обычные протоны, то по рассеянию нейтронов на протонах, характер которого значительно отличается от такового на дейтронах, можно определить расстояния между центрами масс этих двух белков. Усовершенствованный метод нейтронного рассеяния позволяет

определять не только расстояния между рибосомными белками, но и пространственную организацию самих полипептидных цепей в составе рибосомных субчастиц. Полученные таким образом *нейтронные карты* основаны на измерении расстояний между 93 белками. Такие карты имеют фундаментальное значение в интерпретации экспериментальных данных, полученных другими методами, особенно в результате *молекулярного моделирования*.

При отсутствии данных рентгеноструктурного анализа высокого разрешения молекулярные биологи традиционно обращаются к молекулярному моделированию пространственных структур. В некоторых случаях такой подход бывает весьма успешным, что особенно ярко проявилось при расшифровке пространственной структуры ДНК. Все модели, описанные в настоящее время, учитывают филогенетические особенности вторичной структуры 16S рРНК, и в некоторых из них принимаются во внимание третичные взаимодействия внутри этих макромолекул. В последнее время для таких целей все чаще используется компьютерный анализ. Применение вычислительной техники сводит к минимуму субъективизм в построении моделей и позволяет систематически исследовать возможные конформационные состояния анализируемых молекулярных объектов. При этом выбор конкретного алгоритма в современном моделировании оказывает меньшее влияние на конечный результат, чем выбор имеющихся экспериментальных данных и использование ограничивающих условий.

На рис. 1.18 представлена современная модель пространственной структуры 70S рибосомы *E. coli*, разработанная в лаборатории Д. Франка (США) с учетом данных, которые были получены с помощью всех вышеперечисленных методов.

2.4.2. Этапы биосинтеза белка

Хотя построение первых моделей механизмов биосинтеза белка было начато еще в начале 1960-х гг., полное описание процесса трансляции далеко до завершения и в настоящее время. Ниже будут кратко рассмотрены основные черты классической модели биосинтеза белка рибосомами *E. coli*, а также особенности некоторых альтернативных моделей.

Процесс биосинтеза белка рибосомами, как и биосинтез любой другой

макромолекулы клетки, условно разделяют на три основных этапа: инициацию, элонгацию и терминацию. Во время инициации трансляции происходит сборка нативной 70S или 80S рибосомы на транскрибируемой мРНК и подготовка к образованию пептидной связи между первыми двумя N-концевыми аминокислотными остатками синтезируемого полипептида. При элонгации наблюдается последовательное удлинение растущей цепи полипептида аминокислотными остатками, а терминация трансляции сопровождается прекращением синтеза полипептида и его высвобождением из трансляционного комплекса. При этом наблюдается также разделение рибосомы и мРНК, после чего они вступают в новый цикл трансляции. В ходе трансляции рибосома последовательно перемещается вдоль транскрибируемой молекулы мРНК, считывая заключенную в ней генетическую информацию в виде триплетного генетического кода. Трансляция начинается в 5'-концевой части мРНК, а завершается в ее 3'-концевой части. При этом биосинтез полипептида начинается с его N-концевой аминокислоты. Рассмотрим каждый из вышеперечисленных этапов более подробно на примере белоксинтезирующей системы *E. coli*.

Инициация трансляции. Биосинтез белка рибосомами начинается с образования комплекса между малой 30S субчастицей рибосом, инициаторной тРНК и участком транскрибируемой мРНК, содержащим сайт связывания рибосом, который включает в себя иницирующий (как правило, AUG) кодон. В образовании инициационного комплекса с 30S субчастицей принимают участие три белковых фактора инициации – IF1, IF2 и IF3. В ходе этого процесса расходуется одна молекула GTP, которая взаимодействует с IF2 и изменяет его конформацию (см. рис. I.19). Таким образом, на первом этапе образования инициационного комплекса происходит объединение свободной 30S субчастицы с факторами инициации и GTP, после чего с ними последовательно связываются мРНК и инициаторная тРНК (в случае *E. coli*, как правило, формилметионил(fMet)-тРНК^{fMet}). Инициаторная тРНК строго специфична для этой стадии белкового синтеза. Сначала она обычным путем акцептирует Met с образованием Met-тРНК^{fMet}, а затем специальная ферментная система *E. coli* формилирует NH₂-группу остатка Met. Последовательность присоединения инициаторной тРНК и мРНК к 30S субчастице не имеет значения, что и отражено на рис. I.19 (стадии А, А', В,

В').

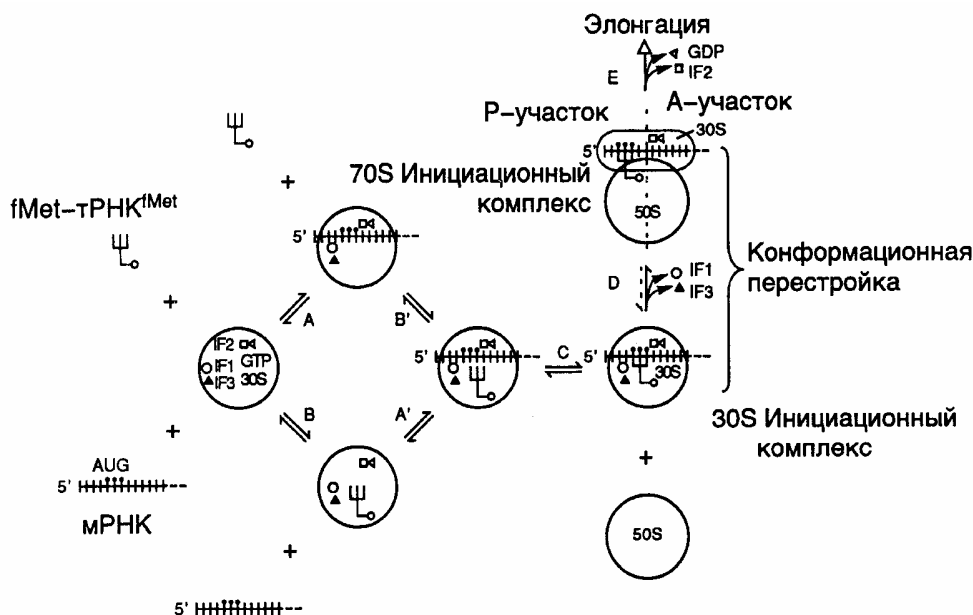


Рис. 1.19. Этапы инициации трансляции у *E. coli*

Вначале после объединения факторов инициации трансляции, GTP, fMet-тРНК^{fMet} и мРНК с 30S субчастицей антикодон тРНК еще не взаимодействует с инициаторным AUG-кодоном (стадии A' и B'). Такое продуктивное взаимодействие тРНК с мРНК происходит в дальнейшем (стадия C), и этот переход является одной из лимитирующих стадий всего процесса образования инициационного комплекса. С завершением стадии C происходит формирование стабильного тройного (из трех основных компонентов) инициационного комплекса, сопровождаемое конформационными перестройками всех его компонентов. После выхода из комплекса факторов инициации трансляции IF1 и IF3 тройной комплекс приобретает способность связывать большую 50S субчастицу рибосом, что сопровождается дальнейшими конформационными перестройками всей рибосомы (стадия D). В ходе этого процесса происходит расщепление молекулы GTP до GDP и ортофосфата и освобождение из комплекса фактора IF2 (стадия E). Формилметионил-тРНК^{fMet} вместе с иницирующим AUG-кодоном перемещаются в донорный (P) участок рибосомы, освобождая акцепторный (A) участок для следующей аминокислотированной тРНК. В результате

инициационный комплекс становится полностью подготовленным для вступления в следующую фазу биосинтеза белка – элонгацию полипептидных цепей.

Элонгация. В соответствии с обсуждаемой моделью принято считать, что после образования тройного комплекса, включающего 70S рибосому, мРНК и инициаторную тРНК, завершается этап инициации трансляции, и процесс биосинтеза белка вступает в фазу элонгации, которая завершается освобождением полипептидных цепей из элонгирующих комплексов. Во время элонгации происходит последовательное присоединение аминокислотных остатков к С-концевым частям строящихся полипептидных цепей, направляемое кодонами транслируемых матричных РНК.

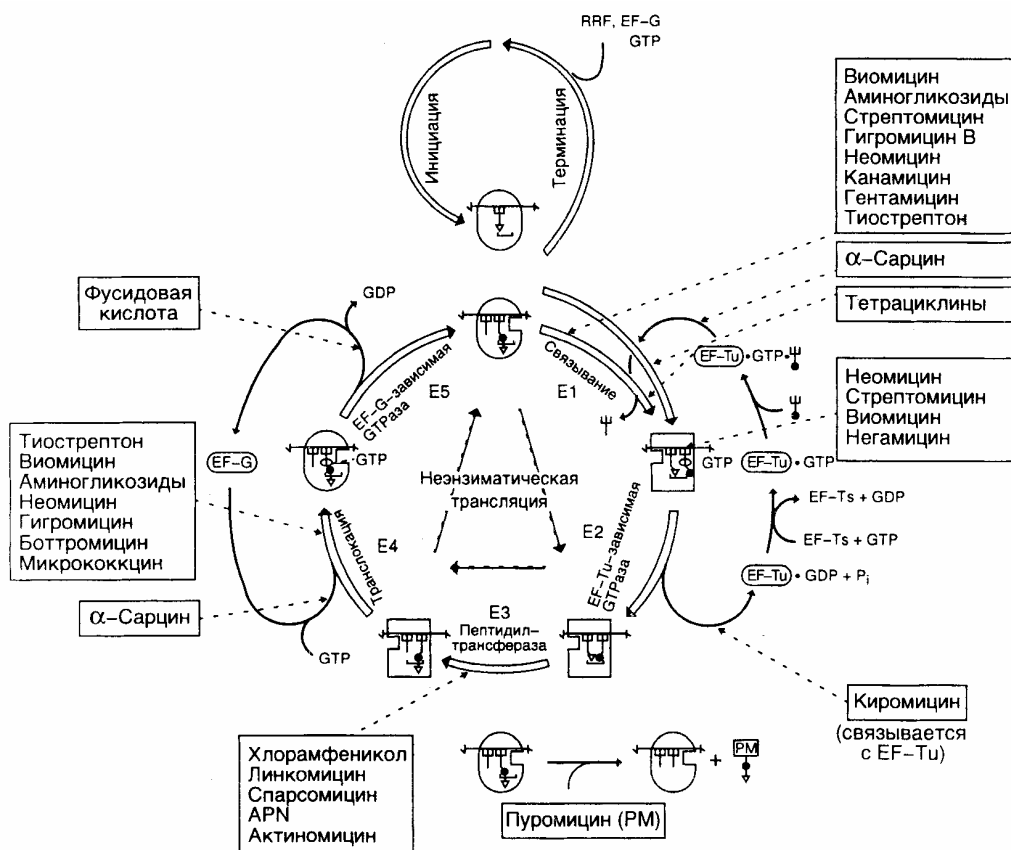


Рис. 1.20. Элонгация полипептидных цепей

Стрелки указывают на реакции, подавляемые соответствующими антибиотиками

Этап элонгации начинается со взаимодействия фактора элонгации

трансляции EF-Tu, молекулы GTP и очередной аминоацилированной тРНК с А-участком рибосомы (см. рис. 1.20, стадия E1). Вхождение аминоацилированной тРНК в А-участок происходит в соответствии с установленным в нем кодоном транскрибируемой мРНК. При этом лишь та аминоацилированная тРНК прочно связывается с рибосомой, у которой антикодон комплементарен кодону, установленному в А-участке. После гидролиза GTP и освобождения EF-Tu•GDP из комплекса (стадия E2) происходит образование новой пептидной связи между карбоксильной группой формилметионина инициаторной тРНК и NH₂-группой аминокислотного остатка, находящегося в А-участке рибосомы в составе аминоацил-тРНК (стадия E3). Эта стадия получила название *транспептидации*. Обмен GDP на GTP в освободившемся комплексе EF-Tu•GDP происходит с участием фактора EF-Ts.

Образовавшийся в итоге пептид удерживается рибосомой через остаток тРНК, находящийся в А-участке, а освободившаяся тРНК временно сохраняется в так называемом Е-участке рибосомы (от англ. exit – выход). Такая соединенная с пептидом тРНК получила название *пептидил-тРНК*. Образовавшаяся пептидил-тРНК далее переносится из А- в Р-участок рибосомы. Эта стадия элонгации (E4) известна под названием *транслокации*. Транслокация индуцируется фактором элонгации EF-G, который освобождается из элонгирующего комплекса после расщепления молекулы GTP. Таким образом, энергия еще одной молекулы GTP используется в акте транслокации. После завершения транслокации происходит освобождение фактора EF-G из элонгирующего комплекса. При этом А-участок рибосомы остается свободным. Следующий цикл элонгации начинается с вхождения в А-участок рибосомы в составе тройного комплекса очередной молекулы тРНК (стадия E1), что сопровождается освобождением формилметионил-тРНК^{fMet} из Е-участка, после чего повторяются все остальные вышеперечисленные стадии элонгации. В физиологических условиях рибосома совершает ~ 20 циклов элонгации в секунду. В соответствии с этим для синтеза белка длиной в 200 аминокислотных остатков требуется ~10 секунд.

В рассмотренной классической модели биосинтеза белка с тремя участками связывания тРНК на любой стадии элонгации с рибосомой взаимодействуют две молекулы тРНК. Иными словами, до стадии транслокации тРНК занимают А- и Р-участки рибосомы, тогда как после транслокации

молекулы ассоциированы с Р- и Е-участками. Между участками А и Е существует аллостерическое взаимодействие, что проявляется в отрицательном кооперативном эффекте связывания молекул тРНК этими участками и означает, что только А- или Е-участки рибосомы могут быть заняты молекулой тРНК, и рибосома не содержит одновременно занятыми оба участка.

Молекулярная мимикрия фактора элонгации EF-G. Недавно было обнаружено, что пространственная структура домена IV полипептидной цепи фактора элонгации EF-G имитирует структуру тРНК в ее комплексе с другим фактором элонгации EF-Tu. При этом структура соответствующей части полипептидной цепи EF-G напоминает форму антикодонной петли тРНК в комплексе с фактором элонгации, ее положение относительно коровой части EF-Tu и даже распределение электростатических зарядов на поверхности полипептида, которое соответствовало таковому углевод-фосфатного остова тРНК. Это открытие позволило по-новому посмотреть на механизм действия EF-G в цикле трансляции. Такого рода данные позволили предположить, что домен IV фактора EF-G занимает во время некоторых этапов транслокации ту же часть А-участка рибосом, что и тРНК. Однако остается непонятным, каким образом это может физически влиять на прохождение акта транслокации.

Терминация трансляции. Процесс трансляции вступает в завершающую фазу после того, как в А-участок рибосомы попадает терминирующий (бессмысленный) кодон мРНК, а пептидил-тРНК перемещается в донорный Р-участок рибосомы. Белковые факторы RF1 и RF2 участвуют в распознавании последовательностей нуклеотидов терминирующего кодона. Фактор RF3, также как и EF-G, имитирует структурные особенности фактора EF-Tu, что дает ему возможность взаимодействовать с А-участком рибосомы. Но поскольку с ним не связана аминоацилированная тРНК, его присутствие в А-участке приводит к обрыву строящейся цепи полипептида. После формирования такого комплекса происходит расщепление сложноэфирной связи между полипептидом и тРНК, а также освобождение синтезированного полипептида. С помощью мутационного анализа было установлено, что молекулы рРНК обеих субчастиц рибосом *E. coli* участвуют в гидролизе пептидил-тРНК.

Для того чтобы рибосома оставшегося комплекса рибосома–мРНК–тРНК могла вступить в следующий цикл трансляции, она должна освободиться из

него. Установлено, что вышеупомянутые рибосомные рилизинг-факторы (RF) совместно с фактором EF-G при участии молекулы GTP обеспечивают диссоциацию комплекса на составные компоненты, которые затем вступают в новый раунд белкового синтеза. Фактор RF4 (иначе называемый RRF – ribosome-recycling factor) не имеет аналога у эукариот. Его роль, по-видимому, заключается в стимуляции перемещения молекулы деацилированной тРНК из Р-участка в Е-участок рибосомы и(или) удалении оставшегося RF-фактора из А-участка. Это способствует полному освобождению рибосомы и ее вовлечению в новый цикл трансляции в результате инициации или реинициации синтеза белка. Отделившаяся от мРНК рибосома перед вступлением в новый цикл диссоциирует на две субчастицы под действием фактора инициации трансляции IF3. Альтернативно, в том случае, если новый иницирующий кодон полицистронной матрицы находится достаточно близко от стоп-кодона, синтез белка может быть реиницирован.

Реинициация трансляции. *Реинициацией* трансляции называют повторное вступление рибосом, терминировавших биосинтез белка, в цикл трансляции без предварительного отделения их от мРНК. Реинициация синтеза белка широко распространена у *E. coli* и играет важную роль в контроле экспрессии генов этого микроорганизма на уровне трансляции. Это связано с тем, что значительная часть бактериальных мРНК полицистронна, а, следовательно, за терминирующим кодоном одного цистрона на небольшом от него удалении располагается иницирующий кодон следующего. Благодаря реинициации имеет место координированная (сопряженная) трансляция нескольких ОРС, объединенных в полицистронной матрице.

Имеются данные о том, что рибосомы *E. coli*, терминировавшие синтез полипептида, обладают способностью перемещаться на короткие расстояния в окрестностях терминирующего кодона и после встречи с иницирующим кодоном в новом сайте инициации трансляции могут начать следующий раунд трансляции без отделения от мРНК. Новый иницирующий кодон может располагаться выше или ниже стоп-кодона предыдущего гена, а может и перекрываться с ним (например, в последовательности AUGA). Реинициация может происходить с полной эффективностью, если в сайте реинициации имеется адекватная SD-последовательность, а терминирующий и иницирующий кодоны расположены достаточно близко друг к другу (менее

эквивалента длины рибосомы). Последнее обстоятельство указывает на быструю кинетику отделения терминировавших рибосом от мРНК.

Важным следствием сопряжения трансляции у прокариот через реинициацию является зависимость экспрессии целой серии генов, принадлежащих одному оперону, от трансляции первой ОРС полицистронной матрицы. При этом следует заметить, что сопряжение экспрессии генов на уровне трансляции может оказывать и негативное влияние на эффективность трансляции нижерасположенных цистронов.

Если обычная инициация трансляции у *E. coli* на кодоне UUG происходит с очень низкой эффективностью, то реиницирующие рибосомы используют его для начала синтеза белка весьма охотно. Это объясняют отсутствием фактора IF3 в реиницирующем комплексе, который во многом определяет точность выбора иницирующего кодона рибосомами.

На поведение рибосом, терминировавших синтез белка, большое влияние оказывают и факторы терминации трансляции (RF). В опытах с мутантным бактериофагом R17, содержащим амбер-кодон в положении 7 гена белка оболочки, было установлено, что в отсутствие фактора терминации трансляции RF4 вслед за терминацией трансляции на бессмысленном кодоне имела место реинициация трансляции на следующем кодоне мРНК, что завершалось синтезом белка оболочки, укороченного с N-конца на 7 аминокислот. На этом основании полагают, что рилизинг-фактор RF4 в обычных условиях предотвращает распознавание аминоацилированной тРНК кодона, находящегося в А-участке рибосомы, который в мРНК следует за терминирующим.

Альтернативные модели цикла трансляции. В рассмотренной выше классической модели трансляции перемещение молекул тРНК на большой и малой субчастицах рибосом сопряжено друг с другом. В отличие от этого, в активно обсуждающейся *модели гибридных состояний* (hybrid states model) перемещение тРНК между А- и Р-участками 30S субчастицы происходит независимо от перемещения тРНК между А-, Р- и Е-участками большой субчастицы. В соответствии с этой моделью аминоацил-тРНК попадает в пептидил-тРНК-рибосомный комплекс в составе тройного комплекса EF-Tu•GTP•тРНК и взаимодействует с ней первоначально в гибридном состоянии А/Е. В этом состоянии антикодоновая часть тРНК связывается с А-участком

30S субчастицы, а ее ССА-конец, удерживаемый EF-Tu, располагается в Е-участке большой субчастицы и, возможно, частично на малой субчастице. Вслед за гидролизом GTP происходит освобождение EF-Tu, что делает возможным перемещение ССА-конца аминоацил-тРНК в А-участок большой субчастицы, приводящее к возникновению А/А-состояния, эквивалентного состоянию взаимодействия аминоацил-тРНК с А-участком в классической модели. После образования пептидной связи аминоацил-тРНК, уже связанная с растущей полипептидной цепью, перемещается в Р-участок большой субчастицы, а деацилированная тРНК переходит в Е-участок большой субчастицы. Вновь образованная пептидил-тРНК находится теперь в гибридном А/Р-состоянии: антикодонная часть остается в А-участке 30S субчастицы, а ССА-конец занимает Р-участок большой субчастицы рибосом. При этом деацилированная тРНК находится в гибридном Р/Е-состоянии: антикодонный конец остается в Р-участке малой субчастицы, тогда как ССА-конец занимает Е-участок большой субчастицы. Далее фактор элонгации EF-G в GTP-зависимой реакции обеспечивает перемещение антикодонной части тРНК, находящейся в гибридном состоянии, вместе с мРНК относительно 30S субчастицы. При этом пептидил-тРНК переходит в чувствительное к пурамицину Р/Р-состояние, соответствующее ее взаимодействию с Р-участком в классической модели, а деацилированная тРНК находится в Е-состоянии и на этом этапе трансляции может взаимодействовать только с Е-участком большой субчастицы рибосом.

По крайней мере, три интересных следствия вытекают из модели гибридного состояния. Во-первых, пептидная часть растущего пептида остается на рибосомах в стационарном состоянии, а во время трансляции перемещается тРНК. Во-вторых, транслокация тРНК происходит в два этапа: во время первой стадии обе молекулы тРНК движутся относительно большой субчастицы, на втором этапе обе молекулы тРНК вместе со связанной с ними мРНК перемещаются относительно малой 30S субчастицы рибосом. В-третьих, в процессе синтеза белка имеют место не два или три состояния связывания тРНК, а шесть или даже, возможно, семь таких состояний.

С помощью физических методов были получены прямые доказательства спонтанного прохождения стадии транслокации, опосредуемой пептидилтрансферазой рибосом. При использовании флуоресцентных зондов, связанных с различными участками тРНК и рибосом, удалось обнаружить

изменения в квантовом выходе флуоресценции и анизотропные эффекты при образовании пептидной связи, что указывало на перемещение молекулы тРНК относительно рибосомных белков S21 и L11. На основании этих данных было высказано предположение, что во время пептидилтрансферазной реакции пептидильная цепь остается в постоянном положении относительно рибосомы, а перемещаются молекулы тРНК. Эта модель пептидилтрансферазной реакции получила название *модели перемещения* (displacement model). Она обладает многими общими чертами с моделью гибридного состояния, однако отличается тем, что в этой модели движение мРНК в пептидилтрансферазной реакции сопровождается перемещением тРНК.

Замечания о точности трансляции. Сама по себе стабильность кодон–антикодоновых взаимодействий не может обеспечивать наблюдаемую высокую точность трансляции. Рибосомы активно участвуют в акте распознавания молекулами тРНК соответствующих кодонов мРНК, повышая точность функционирования этого механизма, по крайней мере, на четыре порядка. Данный эффект объясняют функционированием механизмов, корректирующих ошибки на этом этапе трансляции, которые сопряжены с EF-Tu-зависимым гидролизом GTP во время выбора соответствующей аминоацилированной тРНК. Недавние измерения скорости гидролиза GTP рибосомами в присутствии правильной (cognate) или неправильной (noncognate) тРНК и искусственной матрицы показали, что в первом случае она выше в $\sim 10^4$ раз. Это приводит к преимущественному освобождению комплекса EF-Tu•GDP из рибосом, содержащих правильные аминоацил-тРНК в А-участке. Для реализации данного механизма рибосомы должны распознавать правильные и ошибочные кодон-антикодоновые взаимодействия, а также передавать эту информацию своему GTPазному центру.

Мутантные рибосомы, для которых характерна пониженная точность трансляции, как правило, обладают более высоким сродством к тРНК в А-участке. Напротив, у "сверхточных" рибосом такое сродство понижено. В соответствии с этим повышенную точность трансляции можно объяснять в терминах уменьшения неспецифического связывания аминоацилированных тРНК А-участком рибосом и *vice versa*. Недавно было показано, что сродство тРНК к Р-участку таких мутантных рибосом изменяется на противоположное таковому А-участка: у gam-мутантов с низкой точностью трансляции

наблюдается пониженное сродство Р-участка к тРНК, а у "сверхточных" рибосом это сродство повышено. Таким образом, в настоящее время полагают, что простые реципрокные отношения связывают А- и Р-участки рибосом с механизмами, которые управляют взаимодействием мРНК и соответствующих тРНК с рибосомами.

Новая 10Sa РНК, функционирующая в трансляции. В том случае, если транскрибируемая бактериальная мРНК укорочена в своей 3'-концевой части и рибосома не находит в соответствующей рамке считывания терминирующий кодон, она не может закончить трансляцию с помощью стандартного механизма терминации. Полагают, что, столкнувшись с такой ситуацией, рибосома останавливается на конце мРНК и с ней взаимодействует недавно открытая 10Sa РНК, кодируемая у *E. coli* геном *ssrA*, которая запускает процесс деградации частично синтезированного полипептида. Эта необычная РНК, аминоацилированная остатком Ala на своем 3'-конце, подобном стандартной акцепторной последовательности тРНК, сочетает в себе свойства тРНК и мРНК. 10Sa РНК, как и тРНК^{Ala}, образует внутреннюю пару оснований G3:U70, которая необходима для специфического распознавания тРНК^{Ala} рибосомами. После взаимодействия с рибосомой эта РНК вступает в стандартный цикл трансляции: ее остаток Ala включается в недостроенную цепь полипептида. Вслед за этим рибосома транскрибирует короткую последовательность нуклеотидов 10Sa РНК, которая теперь функционирует в качестве матрицы, добавляя десять аминокислот – ANDENYALAA – в С-конец укороченной полипептидной цепи, и терминирует трансляцию на UAA-кодоне 10Sa РНК. С-Концевая олигопептидная последовательность освобожденного полипептида далее распознается специфической протеиназой, которая расщепляет этот полипептид, обеспечивая его утилизацию бактериальной клеткой в процессе катаболизма.

2.4.3. Антибиотики, действующие на уровне трансляции

На рис. 1.21 приведены некоторые широко распространенные антибиотики, являющиеся ингибиторами биосинтеза белка у бактерий. Многие из них находят применение не только как лекарственные средства, но и как превосходные инструменты исследования механизма различных этапов биосинтеза белка. Биохимический анализ мутантов бактерий, устойчивых к

действию конкретных антибиотиков, позволяет обнаруживать сайты действия антибиотиков на рибосомах и идентифицировать изменения компонентов системы белкового синтеза под влиянием этих мутаций. Как правило, для возникновения устойчивости к антибиотику достаточно замены одного аминокислотного остатка из ~7500 остатков белков, составляющих бактериальную рибосому. То же самое относится и к ~4500 основаниям рибосомных РНК. В этом случае не только замены одиночных оснований в рРНК, но и модификация (метилирование) единственного основания могут приводить к подобным эффектам. Рассмотрим механизм действия некоторых антибиотиков более подробно.

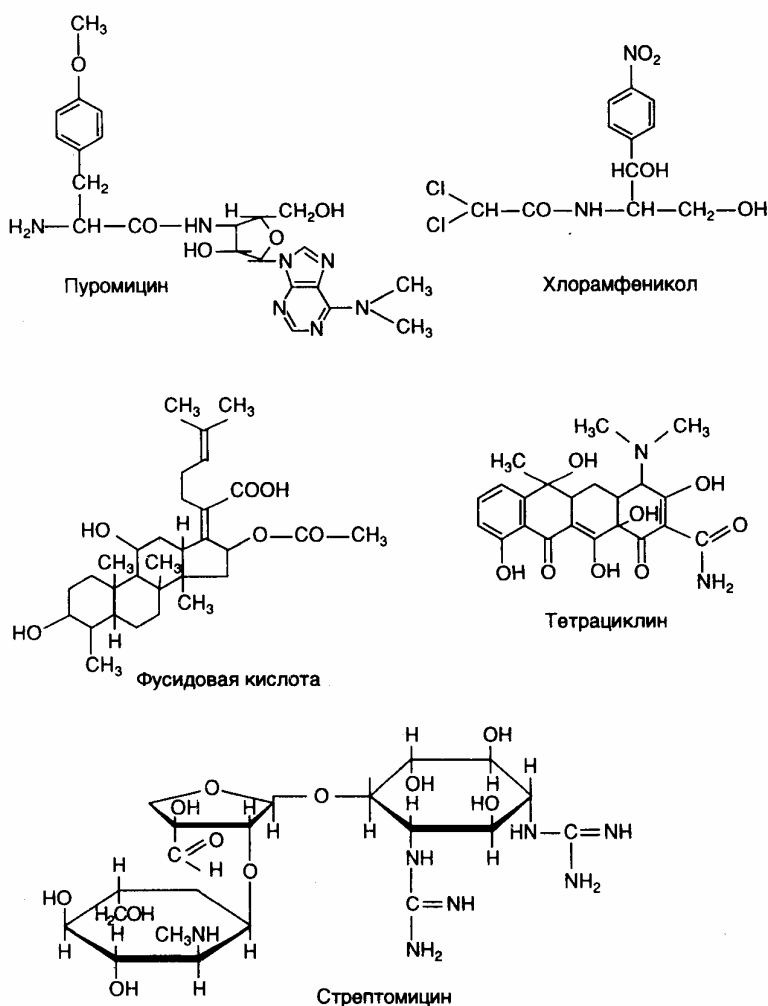


Рис. 1.21. Антибиотики – ингибиторы биосинтеза белка у бактерий

Пурамицин. Этот антибиотик представляет собой производное нуклеозидов и является структурным аналогом 3'-концевой аминокислотированной группировки тРНК. Прямые эксперименты было показано, что пурамицин конкурентным образом замещает очередную аминокислот-тРНК в А-сайте рибосом в процессе трансляции. Он участвует в акте образования пептидной связи в рибосоме, подменяя при этом очередную аминокислот-тРНК. В ходе реакции транспептидации происходит переброска С-конца растущего пептида от пептидил-тРНК на свободную аминогруппу его аминокислотного остатка, что приводит к освобождению пептидил-пурамицина из рибосом и прекращению биосинтеза белка. Пурамицин одинаково хорошо подавляет биосинтез белка как прокариотическими, так и эукариотическими рибосомами.

Хлорамфеникол. Участок лабильного связывания этого антибиотика локализован на 50S субчастице рибосом. Хлорамфеникол полностью ингибирует реакцию пурамицина с пептидил-тРНК, выступая его конкурентным ингибитором. При этом синтез пептида полностью прекращается, и он остается связанным с рибосомами. Предполагают, что хлорамфеникол имитирует аминокислотный конец молекулы аминокислот-тРНК, а его дихлорацетамидная группировка соответствует аминокислоте. Местом действия хлорамфеникола является А-участок 50S субчастицы рибосом, где антибиотик конкурирует с аминокислотным концом молекулы аминокислот-тРНК, препятствуя ее вхождению в А-участок, что сопровождается подавлением биосинтеза белка. В отличие от пурамицина хлорамфеникол ингибирует только бактериальные рибосомы. Сходным механизмом действия обладают антибиотики *линкомицин* и *спарсомицин*. Последний делает ассоциацию пептидил-тРНК с Р-участком рибосом более прочной. При этом хлорамфеникол и линкомицин способны вытеснять спарсомицин из его комплекса с рибосомами.

Фузидовая кислота – антибиотик стероидной природы, блокирует биосинтез белка на стадии транслокации. Его мишенью является не столько сама рибосома, сколько белковый фактор EF2(EF-G), который, как указывалось выше, необходим для GTP-зависимой транслокации. Фузидовая кислота не влияет на взаимодействие фактора EF2(EF-G) и GTP с пре-транслоцированной рибосомой и последующее расщепление GTP. По-видимому, антибиотик препятствует диссоциации указанного комплекса и сопряженной с ней

транслокации. По тому же механизму фусидовая кислота подавляет трансляцию эукариотическими рибосомами.

Тетрациклины. Антибиотики тетрациклинового ряда специфически связываются с 30S субчастицей рибосом, подавляя реакцию аминоацил-тРНК с рибосомами и свободными 30S субчастицами в присутствии матрицы, но не нарушая связывание самого матричного полинуклеотида. Предполагают, что тетрациклины взаимодействуют с акцепторным тРНК-связывающим участком 30S субчастицы рибосом.

Стрептомицин и другие аминогликозидные антибиотики. Стрептомицин (антибиотик углеводной природы) специфически взаимодействует с определенным структурным белком 30S субчастицы рибосом, блокируя стадию инициации трансляции. В присутствии стрептомицина наблюдается стимуляция связывания аминоацил-тРНК, не соответствующих кодам мРНК, находящимся в данный момент в акцепторном А-участке рибосом. В итоге происходит ошибочное включение аминокислот в полипептидные цепи синтезируемых белков. Это может проявляться в фенотипической супрессии нонсенс-мутаций у мутантных бактерий. Аминогликозидные антибиотики также вызывают неспецифическое связывание матричных полинуклеотидов рибосомами. Следствием является, например трансляция одноцепочечных ДНК рибосомами в бесклеточных системах в присутствии аминогликозидов.

2.5. Трансляция у эукариот

Бактерии обладают единственной универсальной системой трансляции, основные механизмы функционирования которой были кратко рассмотрены выше. В отличие от этого, клетки животных кроме основной системы трансляции, локализованной в цитоплазме, имеют дополнительную систему трансляции митохондрий, которая по ряду свойств приближается к бактериальной. Клетки растений обладают еще одной дополнительной системой биосинтеза белка, функционирующей в хлоропластах. Большинство данных о механизмах биосинтеза белка у эукариот было получено с использованием бесклеточных белоксинтезирующих систем (подробнее о принципах функционирования таких систем см. в разделе 7.4). В последнее время важные результаты о механизмах трансляции у эукариот были получены

с использованием стабильно трансформированных клеток животных и растений, выращиваемых в культуре. В ходе этих исследований установлено, что у растений и животных в основном функционируют одни и те же механизмы трансляции. Ниже будут рассмотрены основные молекулярные механизмы, участвующие в трансляции мРНК у эукариот, с привлечением данных, полученных главным образом на дрожжах *S. cerevisiae*.

2.5.1. Особенности первичной структуры эукариотических мРНК

Зрелая мРНК эукариот наряду с основной последовательностью нуклеотидов, в которой закодирована информация о последовательности аминокислот в соответствующем белке, содержит целый ряд некодирующих последовательностей, присутствие которых исключительно важно для ее эффективной, регулируемой трансляции рибосомами. Одни из этих последовательностей, такие как кэп-группа и 3'-концевая поли(А), не кодируются непосредственно генами, а, как это подробно рассматривалось выше, добавляются ко- и посттранскрипционно. Другие некодирующие последовательности, в том случае, если они не являются продуктами посттранскрипционного редактирования мРНК, имеют генное происхождение. Эти последовательности часто содержат регуляторные сигналы, обеспечивающие определенный уровень трансляции мРНК рибосомами.

Участок мРНК, расположенный между кэп-группой и первым иницирующим кодоном основной открытой рамки считывания (ОРС), которая и несет информацию о последовательности аминокислот в белке, получил название *5'-концевой нетранслируемой области* (5'UTR – 5' untranslated region), или *лидерной последовательности*. Сегмент мРНК, расположенный между последним терминирующим кодоном основной ОРС и началом поли(А)-последовательности, называют *3'-концевой нетранслируемой областью* (3'UTR). Первое название не совсем удачно. Последовательности 5'UTR, как правило, способны образовывать сложные вторичные структуры типа "стебель-петля" и содержать короткие ОРС (uORF – upstream open reading frame), которые оказывают сильное влияние на эффективность трансляции мРНК (см. ниже). Помимо этого 5'UTR могут включать в себя регуляторные последовательности, распознаваемые *транс-действующими* белковыми факторами. Последовательности 5'UTR обеспечивают регулирующую

трансляцию мРНК (и координированную экспрессию соответствующих генов) в онтогенезе многоклеточных организмов.

3'UTR и поли(A)-последовательность оказывают влияние на состояние рибосом после терминации синтеза полипептидных цепей. Кроме того, по крайней мере, 3'-концевая поли(A)-последовательность участвует в инициации трансляции.

2.5.2. Инициация биосинтеза белка эукариотическими рибосомами

Как будет видно из дальнейшего изложения, инициация трансляции эукариотических мРНК может осуществляться, по крайней мере, тремя способами. В соответствии с первым наиболее распространенным механизмом (модель *сканирования*) рибосомы после взаимодействия с 5'-концевой последовательностью мРНК осуществляют поиск иницирующего AUG-кодона, перемещаясь вдоль 5'UTR. При реализации второго механизма рибосомы иницируют биосинтез белка на *внутренних* AUG-кодонах, удаленных от 5'-концевой кэп-группы. И, наконец, после освобождения полипептида из транслирующего комплекса рибосомы, не отделяясь от мРНК, способны *реиницировать* биосинтез белка на следующем иницирующем кодоне.

Факторы инициации трансляции. Большинство молекулярных механизмов, осуществляющих регуляцию экспрессии генов на уровне трансляции, реализуется на стадии инициации биосинтеза белка. По-видимому, этот факт находит свое отражение в большой сложности аппарата инициации трансляции. Помимо субъединиц эукариотических рибосом и белков, обычно ассоциированных с 5'- и 3'-концевыми последовательностями мРНК, в инициации принимают участие по меньшей мере 11 белковых факторов, построенных более чем из 25 полипептидов (табл. I.11).

Таблица I.11

Факторы инициации трансляции дрожжей *S. cerevisiae*

Фактор	Субъединица	Предполагаемая функция
eIF1		Обеспечивает связывание Met-тРНК и мРНК с 40S субчастицей рибосом и распознавание иницирующего AUG-кодона (у животных)

eIF1A		Обеспечивает диссоциацию 40S–60S субчастиц рибосом, связывание Met-тРНК и распознавание иницирующего кодона (у животных)
eIF2	α, β, γ	Участвует в выборе иницирующего кодона
eIF2B	α, β, γ, δ, ε	Обеспечивает обмен гуанилового нуклеотида на eIF2
eIF3	α, β, γ, δ, ε, ζ, η, θ	Обеспечивает связывание Met-тРНК и мРНК с 40S субчастицей и диссоциацию 40S–60S субчастиц рибосом
eIF4A		АТРаза, РНК-связывающая хеликаза
eIF4B		Хеликаза, облегчающая связывание РНК
eIF4E		Взаимодействует с кэп-группой мРНК
eIF4G	G1(p150), G2(p130)	Взаимодействует с eIF3, eIF4E и Pab 1p
eIF4H		Стимулирует активность eIF4B и компонентов eIF4F
eIF5		Вызывает диссоциацию факторов инициации
eIF5A		Функции неизвестны, мутации изменяют стабильность мРНК
eIF6		Вызывает диссоциацию 40S–60S субчастиц рибосом

Учитывая сложность процесса инициации трансляции у эукариот, последовательность реакций, приводящих к образованию первой пептидной связи в строящемся полипептиде, удобно разбить на ряд последовательных этапов, что является сознательным упрощением единого процесса.

Взаимодействие мРНК с кэп-связывающим комплексом и рибосомами. Возможность вступления эукариотических мРНК в цикл трансляции как правило обеспечивается их 5'-концевыми кэп-структурами, с которыми взаимодействуют белки кэп-связывающего комплекса (СВС). Хотя основными компонентами СВС являются факторы инициации трансляции, его роль далеко не ограничивается участием в инициации синтеза белка рибосомами. Как уже обсуждалось в разделе 2.2.4, полифункциональные белки СВС интегрируют основные реакции метаболизма мРНК и их предшественников в эукариотических клетках, необходимые для осуществления эффективной регулируемой трансляции.

Взаимодействие eIF2 с Met-мРНК. Гетеротримерный фактор eIF2 обеспечивает взаимодействие рибосом с инициаторной Met-тРНК и мРНК *in vitro*. Гены всех трех субъединиц являются жизненно важными. В связанном с GTP состоянии eIF2 приобретает способность взаимодействовать с Met-тРНК с образованием тройного комплекса Met-тРНК–eIF2–GTP. Имеются данные, указывающие на участие фактора eIF2B в обмене GDP на GTP в комплексе eIF2–GDP. Неизвестна точная последовательность объединения тройного комплекса с рибосомой и мРНК. Большая часть имеющихся данных указывает на то, что взаимодействие тройного комплекса с 40S субчастицей рибосом предшествует образованию комплекса 40S–мРНК. Однако наличие феномена реинициации трансляции, при которой тройной комплекс входит в иницирующий комплекс с предсуществующим комплексом рибосома–мРНК, указывает на возможность осуществления этих событий в другой последовательности.

Формирование кэп-связывающего комплекса на мРНК. Сборка прединициационного комплекса на мРНК начинается со взаимодействия фактора eIF4E с кэп-группой мРНК. Это дает возможность объединения eIF4E и eIF4G с образованием многокомпонентного фактора eIF4F. В настоящее время не исключается возможность того, что объединение eIF4E и eIF4G предшествует взаимодействию первого с кэп-группой. У животных в состав многокомпонентного фактора eIF4F, кроме того, входит eIF4A, причем eIF4G удерживает два других фактора рядом друг с другом. Комплекс факторов eIF4F животных обладает двунаправленной АТФ-зависимой РНК-хеликазной активностью, которая стимулируется фактором eIF4B. Недавно (1998 г.) обнаруженный фактор eIF4H усиливает активность eIF4F и eIF4B, однако его истинная роль в инициации трансляции остается невыясненной. Еще два белка дрожжей взаимодействуют с компонентами eIF4F: белок p20, конкурирующий с eIF4G за связывание eIF4E, а также поли(А)-связывающий белок Pab 1p, который контактирует со специфическим сайтом полипептидной цепи фактора eIF4G1. Функциональным аналогом p20 у млекопитающих является белок 4E-BP – ингибитор инициации трансляции.

Полипептидная цепь eIF4G млекопитающих содержит сайты связывания факторов eIF3, eIF4E и eIF4A. На этом основании делается вывод, что eIF4G в обеих системах выполняет функции белка-адаптера, обеспечивающего сборку

комплекса eIF4F. Исключительно важная роль фактора eIF4E в регуляции экспрессии генов на уровне трансляции (и в канцерогенезе) будет рассмотрена в разделе 3.4.1.

Таким образом, взаимодействию малой субчастицы рибосом с мРНК предшествует серия высоко специфических белок–белковых и белково–нуклеиновых взаимодействий, приводящих к формированию белкового комплекса вокруг кэп-группы мРНК, в котором полипептидная цепь фактора eIF4G обладает сайтом связывания eIF3. Последний, в свою очередь, специфически взаимодействует с малой субчастицей рибосом, обеспечивая ее вхождение в прединициационный комплекс. При этом 40S субчастица ассоциирована с тройным комплексом Met-тРНК–eIF2–GTP, содержащим аминокислотилированную инициаторную тРНК^{Met}.

В итоге образуется прединициационный комплекс, содержащий мРНК, 40S субчастицу рибосом, связанную с тройным комплексом Met-тРНК–eIF2–GTP и через фактор eIF3 взаимодействующую с фактором eIF4G. Последний, в свою очередь, является частью многокомпонентного фактора eIF4F, в который кроме eIF4G входят eIF4E, взаимодействующий с кэп-группой мРНК, и eIF4A, обладающий РНК-хеликазной активностью. Кроме того, в состав этого комплекса входит фактор eIF1A. В таком виде прединициационный комплекс способен перемещаться вдоль 5'UTR мРНК и осуществлять поиск иницирующего AUG-кодона.

Роль 3'-концевой поли(А)-последовательности мРНК в инициации трансляции. Помимо вышеупомянутых сайтов белок–белковых взаимодействий, N-концевая часть полипептидной цепи дрожжевого eIF4G содержит участок, взаимодействующий с поли(А)-связывающим белком Pab1p. Другим указанием на участие 3'-концевой поли(А)-последовательности мРНК в трансляции является наличие мутаций в генах рибосомных белков 60S субчастицы, супрессирующих мутации в гене *pab1*. Кроме того, как будет видно из дальнейшего изложения, 3'-концевая поли(А)-последовательность мРНК может обеспечивать кэп-независимую инициацию трансляции. Данные такого рода указывают на возможную ключевую роль этой последовательности в инициации синтеза белка, однако механизм данного явления остается неизвестным.

Выбор точки инициации трансляции и инициация биосинтеза белка.

Сформировавшись, прединициационный комплекс должен оказаться на иницирующем AUG-кодоне мРНК, в ряде случаев весьма удаленном от кэп-группы, с которой он первоначально взаимодействует. Рибосомы прокариот локализуют точку инициации биосинтеза белка путем непосредственного взаимодействия регуляторных элементов 5'UTR их мРНК (таких, как SD-последовательность), расположенных в области инициации трансляции TIR (translation initiation region), с 3'-концевой последовательностью 16S рРНК малой субчастицы рибосом. В эукариотической клетке не обнаружено подобных взаимодействий между мРНК и рРНК. Одной из наиболее популярных в настоящее время моделей поиска эукариотической рибосомой точки инициации трансляции является модель сканирования, в соответствии с которой прединициационный комплекс перемещается вдоль 5'UTR до первого специфически распознаваемого им иницирующего кодона.

Модель сканирующей рибосомы. Как следует из модели сканирования, сформированный прединициационный комплекс перемещается от кэп-группы мРНК вдоль 5'UTR, "проверяя" ее последовательность на наличие иницирующего AUG-кодона. В настоящее время отсутствуют твердые доказательства того, что сканирование является строго однонаправленным. Из-за отсутствия в мРНК эукариот SD-подобных последовательностей и соответствующих контактов с рРНК AUG-кодон распознается в результате кодон–антикодового взаимодействия с участием Met-тРНК, входящей в состав прединициационного комплекса. Перемещение комплекса часто должно происходить на фоне ярко выраженной вторичной структуры лидерной последовательности мРНК. В этой связи предполагается, что происходящее во время сканирования расщепление ATP сопряжено с работой РНК-хеликазы, разрушающей вторичную структуру 5'UTR. Как уже упоминалось, данная активность ассоциирована с фактором eIF4A, входящим в состав прединициационного комплекса, и стимулируется фактором eIF4B. Однако роль этих факторов, по-видимому, не ограничивается разрушением вторичной структуры лидера, поскольку их присутствие требуется и для инициации синтеза белка на мРНК с короткими 5'UTR, не обладающими выраженной вторичной структурой.

В выборе AUG-кодона у эукариот участвует фактор eIF2. На это указывает тот факт, что его мутационные повреждения сопровождаются ослаблением специфичности такого выбора. Неожиданными оказались недавно полученные результаты, подчеркивающие важную роль фактора eIF5 в распознавании иницирующего кодона прединициационным комплексом. Не исключено, что этот фактор определяет точность процесса распознавания AUG и является функциональным аналогом прокариотического фактора IF3. Описаны мутантные формы eIF5, в присутствии которых *in vivo* в качестве иницирующего узнается кодон UUG. Не исключено, что совместное действие факторов eIF2 и eIF5 в обеспечении точности выбора иницирующего кодона становится возможным благодаря наличию на β -субъединице eIF2 сайта связывания eIF5.

После локализации иницирующего кодона 40S субчастицей она приобретает способность объединяться с большой 60S субчастицей рибосом при участии фактора eIF5, что в конечном итоге приводит к образованию полноценного инициационного комплекса. В это время происходит отделение от комплекса ряда факторов инициации трансляции, сопряженное с гидролизом GTP. Прежде всего, освобождается комплекс eIF2–GDP, а также большинство остальных факторов инициации, включая eIF1A и eIF3. В таком виде при наличии соответствующей аминоацил-тРНК инициационный комплекс способен образовывать первую пептидную связь в строящейся полипептидной цепи, т.е. иницировать синтез белка.

Контекст и приоритеты иницирующих AUG-кодонов.

Последовательности, окружающие иницирующие AUG-кодоны, оказывают сильное влияние на эффективность инициации трансляции у позвоночных и в значительно меньшей степени у дрожжей. В последнем случае наиболее благоприятным для инициации трансляции является нуклеотид А в положении –3 по отношению к AUG (нуклеотид А в AUG-кодоне находится в положении +1), замена которого на любой другой нуклеотид снижает эффективность инициации приблизительно в два раза. Вообще, А-богатые последовательности, предшествующие AUG, характерны для мРНК дрожжей, что отличает их от мРНК позвоночных, соответствующие области мРНК которых сильнее обогащены GC. Это объясняют высокой чувствительностью аппарата трансляции дрожжей к вторичной структуре 5'UTR их мРНК, которая

при наличии GC-пар была бы более прочной. В следующих за AUG последовательностях мРНК дрожжей не обнаружено предпочтения в отношении А, и они, как правило, обогащены пиримидиновыми нуклеотидами.

Оптимальный контекст для инициации трансляции в клетках животных и растений, по-видимому, один и тот же: AACAATGGC. Самыми важными для инициации в обоих случаях являются пурин в положении –3 и G в положении +4. В клетках животных на эффективность трансляции мРНК оказывают влияние также нуклеотиды в положениях +5 и +6. Иницирующие кодоны в контексте, отличающемся от оптимального, узнаются рибосомами менее эффективно и допускают их прохождение до следующего иницирующего кодона. Это явление, обнаруженное в клетках животных и растений, получило название *ослабленного сканирования* (leaky scanning). Некоторые примеры реализации механизма ослабленного сканирования будут рассмотрены в разделе 3.4.1 (см. рис. 1.40,а–в) в связи с особенностями инициации трансляции у вирусов растений.

Другим важным фактором, определяющим выбор AUG-кодона в качестве иницирующего, является его положение в 5'UTR. Как правило, ближайший к 5'-концу мРНК AUG предпочтительно используется для инициации трансляции. Это объясняют преимущественным перемещением сканирующей рибосомы в направлении 5'→3'. Многие 5'UTR содержат дополнительные AUG, не принадлежащие к основным OPC, а также короткие OPC (uAUG, uORF) перед основным иницирующим кодоном. И те и другие обычно оказывают ингибирующее действие на трансляцию соответствующих мРНК. Ингибирующий эффект является наиболее сильным, если расстояние первого uAUG от 5'-конца мРНК меньше 15–20 нуклеотидов. Наличие двух следующих друг за другом AUG-кодонов может сопровождаться их использованием в качестве альтернативных сайтов инициации трансляции. В этом случае одна и та же мРНК может направлять синтез двух полипептидов, различающихся лидерными пептидами, что, в свою очередь, может определять направление их внутриклеточного транспорта.

Распознавание иницирующих кодонов в процессе инициации трансляции может сопровождаться продолжительными паузами в дальнейшем перемещении рибосом вдоль мРНК при ее сканировании.

Иницирующие кодоны, отличающиеся от AUG. Трансляция в

клетках млекопитающих и насекомых может начинаться на кодонах, которые отличаются от канонического AUG. В природных мРНК в качестве альтернативного иницирующего кодона чаще всего встречается CUG и значительно реже – AUC и ACG. В клетках дрожжей любые не-AUG-кодона распознаются очень неэффективно. Эффективность может быть повышена мутациями в субъединицах фактора eIF2 β и γ . Кодон AUU обычно открывает первую ОРС у некоторых вирусов растений. В этом случае эффективность инициации трансляции составляет ~ 10% от эффективности на каноническом кодоне AUG, который располагается ниже первого. Таким образом, использование вирусами на одной матрице неканонического и канонического кодонов является одним из регуляторных механизмов, контролирующих соотношение синтезирующихся полипептидных цепей на уровне трансляции позволяющих рибосоме достичь в процессе сканирования второго иницирующего кодона и иницировать на нем синтез белка.

У млекопитающих вышерасположенный неканонический иницирующий кодон также, как правило, сопровождается каноническим кодоном. При этом дополнительная инициация трансляции на неканоническом кодоне чаще всего характерна для ОРС, кодирующих регуляторные белки. Образование по такому механизму полипептидных цепей, удлиненных с N-концевой части, приводит к появлению у них новой регуляторной активности, а сам процесс инициации на неканоническом кодоне может контролироваться условиями внутри клетки.

Влияние вторичной структуры мРНК на инициацию трансляции.

Вторичные структуры в 5'UTR мешают сканированию мРНК 40S субчастицами рибосом. Уровень ингибирования инициации трансляции зависит от положения таких структур относительно 5'-конца матрицы. Например, в лизатах ретикулоцитов кроликов структура типа стебель–петля со стабильностью –30 ккал/моль ингибирует инициацию трансляции, если расположена в непосредственной близости от кэп-структуры и не замечается рибосомами при удалении от 5'-конца более чем на 52 нуклеотида. В последнем случае у сканирующего прединициационного комплекса хватает энергии для разрушения этой структуры 5'UTR, но он не может быть собран при наличии шпильки, расположенной в непосредственной близости от кэп-структуры. Тем не менее, шпильки, удаленные от 5'-конца мРНК, могут подавлять трансляцию более чем на 85%, если их стабильность превышает –50 ккал/моль.

Интересно отметить, что структуры типа стебель–петля, образованные вслед за иницирующим AUG-кодоном, могут повышать эффективность инициации трансляции. Полагают, что такие структуры вызывают задержку сканирующего комплекса на иницирующем кодоне, что повышает вероятность его правильного распознавания. Например, элемент вторичной структуры, локализованный на 14 нуклеотидов ниже иницирующего кодона, оказывает стимулирующее влияние на распознавание слабых AUG-кодонов, а также иницирующих кодонов, отличающихся от AUG.

Шунтирование при сканировании мРНК. У некоторых вирусов растений, в частности при трансляции 35S РНК вируса мозаики цветной капусты (CaMV), содержащей лидерную последовательность, в процессе инициации имеет место модифицированное сканирование, получившее название *рибосомного шунта* (см. рис.1.40,а). По этому механизму сканирующие рибосомы могут непосредственно переноситься из донорного в акцепторный сайт на РНК, избегая линейного сканирования расположенной между ними последовательности нуклеотидов. Такое явление было исследовано на искусственных матрицах, содержащих между донорным и акцепторным сайтами последовательности нуклеотидов, образующих интенсивную вторичную структуру. При этом в отсутствие функционирования шунта последовательности обладали ингибирующим действием, точно предсказываемым на основании модели сканирования. Интересно, что шунтирование внутренних последовательностей сканирующими рибосомами может происходить, хотя и с меньшей эффективностью, и *in trans*, т.е. в условиях, когда донорный и акцепторный сайты располагаются на разных молекулах РНК. Механизмы, лежащие в основе шунтирования, в настоящее время до конца не изучены. Полагают, что ключевую роль в этом могут играть специфические взаимодействия участков РНК, расположенных на больших расстояниях друг от друга. Механизм шунтирования обнаружен при трансляции РНК вируса Сендай, тройной лидерной последовательности аденовирусов, а также у паповавирусов.

Реинициация трансляции. В отличие от прокариотических, мРНК эукариот, как правило, моностронны, т.е. содержат только одну основную ОРС. Однако как уже упоминалось выше, часто 5'UTR эукариотических мРНК содержат короткие ОРС. Например, у дрожжей известно несколько сотен видов

таких мРНК, и они содержат от одной до шести коротких ОРС, часть из которых может перекрываться. То же самое характерно для мРНК животных. Несмотря на свои малые размеры, короткие ОРС способны обеспечивать инициацию, терминацию и реинициацию трансляции рибосомами эукариот.

Короткие ОРС эукариот играют важную роль в регуляции экспрессии генов на уровне трансляции, о чем подробнее будет говориться в разделе 3.4. Здесь же хочется отметить, что по функциональному признаку короткие ОРС разделяются на две группы: в одних случаях регуляторные функции коротких ОРС не зависят от их кодирующего потенциала, тогда как другие ОРС реализуют свои регуляторные возможности через кодируемые ими пептиды.

В том случае, если короткие ОРС следуют в мРНК друг за другом, рибосомы после завершения трансляции одной из них могут реиницировать синтез белка на следующем иницирующем кодоне. Как и в случае прокариотической трансляции, эффективность реинициации у эукариот в большой степени зависит от расстояния между терминирующим и иницирующим кодоном, а также от нуклеотидного контекста, в котором находится иницирующий кодон.

Инициация на внутренних AUG-кодонах. Третья возможность инициации трансляции у эукариот основана на прямом вхождении рибосом в цикл трансляции независимо от кэп-структур и лидерных последовательностей мРНК при прямом участии их IRES-последовательностей (internal ribosome entry site). В отличие от классической инициации трансляции у прокариот на внутренних AUG-кодонах полицистронных мРНК, обеспечиваемой последовательностями TIR, IRES-опосредованная инициация происходит не на полицистронных мРНК в обычном смысле, т.е. содержащих последовательности нескольких структурных генов. В этой связи полагают, что основная физиологическая роль инициации трансляции на внутренних AUG-кодонах эукариот заключается в предоставлении некоторым ОРС возможности транслироваться, минуя основной кэп-зависимый механизм инициации. Например, во время вирусной инфекции протеиназа пикорнавирусов отщепляет от полипептидной цепи клеточного фактора eIFG N-концевой пептид, содержащий сайт связывания фактора eIF4E – основного кэп-связывающего белка эукариот. При этом С-концевая часть такого модифицированного фактора сохраняет сайты связывания факторов eIF3 и eIF4A. В результате

происходит (неполное) переключение с кэп-зависимой трансляции мРНК клетки-хозяина на кэп-независимую трансляцию вирусных мРНК, опосредованную IRES-последовательностями вирусов.

2.5.3. Элонгация полипептидных цепей

Элонгация полипептидных цепей в ходе эукариотической трансляции традиционно пользовалась меньшим вниманием исследователей по сравнению с инициацией, поскольку считалось, что ее механизмы в основных чертах идентичны таковым бактерий. Дальнейшие исследования показали, что данная точка зрения в основном соответствует действительности, хотя эукариотическая система трансляции обладает более сложным набором факторов элонгации.

Факторы и механизмы элонгации. Эукариотические клетки содержат в большом количестве фактор элонгации eEF1A, который является функциональным гомологом бактериального фактора EF1A(EF-Tu). Так же как и у бактерий, этот фактор образует тройной комплекс с GTP и аминоксил-тРНК, обеспечивая вхождение последней в А-участок элонгирующей рибосомы. Два других эукариотических фактора eEF1B и eEF2 резко отличаются от бактериальных функциональных аналогов EF1B(EF-Ts) и EF2(EF-G) по аминокислотным последовательностям. Гетеротримерный фактор eEF1B, как и его бактериальный аналог, катализирует обмен GDP на GTP в комплексе eEF1A–GDP. Фактор eEF2, по аналогии с бактериальными системами, обеспечивает транслокацию пептидил-тРНК в Р-участок рибосом и перенос деацилированной тРНК в Е-участок. У высших организмов этот фактор служит мишенью регуляторных воздействий через фосфорилирование (см. раздел 3.4). Замечательным свойством факторов eEF1A и eEF2 является способность связываться с компонентами цитоскелета эукариотических клеток. Полагают, что это их свойство может обеспечивать один из механизмов внутриклеточного транспорта мРНК, направляющих ее в полисомы.

Растущий полипептид выводится в цитоплазму через канал, начало которого расположено на поверхности рибосомы, где он взаимодействует с белками, распознающими сигнальную последовательность, или с другими цитоплазматическими факторами, которые обеспечивают его направленный транспорт внутри эукариотической клетки. У бактерий растущая полипептидная

цепь может вызывать уменьшение скорости элонгации, а природа предпоследней аминокислоты оказывает сильное влияние на терминацию трансляции. Предполагают, что такого рода эффекты являются следствием взаимодействия между строящимся пептидом и факторами трансляции, рРНК или непосредственно каналом, через который он переносится к поверхности рибосомы. Подобные механизмы, по-видимому, функционируют и у эукариот. У дрожжей, как и у *E. coli*, скорость элонгации трансляции снижается в присутствии редко встречающихся кодонов в мРНК. Наличие определенного числа таких кодонов вблизи сайта инициации трансляции значительно снижает скорость считывания соответствующих ОРС. На скорость декодирования мРНК рибосомами оказывают влияние и характер фолдинга строящихся полипептидных цепей (см. раздел 3.6.1), а также сигнальные последовательности аминокислот, определяющие направление их посттрансляционного транспорта внутри эукариотических клеток.

Уникальный фактор элонгации eEF-3 дрожжей. Клетки дрожжей и других грибов обладают дополнительным фактором элонгации eEF3, аналог которого пока не обнаружен у животных. Фактор является мономерным белком с молекулярной массой 116 кДа, обладающим АТРазной активностью. Функциональная роль этого фактора заключается в стимуляции вхождения тройного комплекса eEF1A•aa-тРНК•GTP в А-участок элонгирующей рибосомы, что является следствием его стимулирующего влияния на освобождение деацилированной тРНК из Е-участка рибосомы. Освобождение деацилированной тРНК сопряжено с гидролизом АТФ, катализируемым этим фактором. Несмотря на то что рибосомы животных обладают АТРазной активностью, гидролиз АТФ для трансляции, осуществляемой этими рибосомами, не требуется.

Благодаря своей уникальности, фактор eEF3 активно исследуется фармацевтическими компаниями как потенциальная мишень для противогрибковых препаратов.

2.5.4. Терминация трансляции

В эукариотических белоксинтезирующих системах терминация трансляции, как и у бактерий, контролируется специфическими рилизинг-

факторами. Однако у эукариот эти факторы менее разнообразны. В частности, у них отсутствует функциональный аналог бактериального фактора RRF/RF4.

Факторы терминации. По современным представлениям, элонгирующая эукариотическая рибосома распознает стоп-кодоны, находящиеся в одной рамке с основными ОПС, после взаимодействия с гетеродимерным комплексом релизинг-факторов (RF), в состав которого входят факторы eRF1 и eRF3. Фактор eRF1 необходим для распознавания всех трех терминирующих кодонов (UAA, UAG и UGA) и освобождения синтезированного полипептида. Фактор eRF3 является GTPазой, обладающей гомологией с eEF1A, которая, гидролизуя GTP, стимулирует терминацию независимо от последовательности нуклеотидов в терминирующих кодонах.

Влияние нуклеотидного контекста на эффективность терминации. Два основных фактора оказывают влияние на эффективность терминации трансляции у эукариот. Этими факторами являются последовательности нуклеотидов в окрестностях терминирующих кодонов и структура С-концевой части строящейся полипептидной цепи. Терминирующие кодоны дрожжей по частоте их использования можно расположить в следующий ряд: UAA(53%) > UGA(27%) > UAG(20%). Если анализировать только активно экспрессирующиеся гены, то частота использования UAA оказывается еще большей – 87%. Анализ последовательностей нуклеотидов в окрестностях терминирующих кодонов показал, что и они не являются случайными. Путем исследования способности следующего за стоп-кодом нуклеотида изменять эффективность его супрессии в гене *lacZ* было установлено их следующее влияние на усиление терминации трансляции на соответствующих терминирующих кодонах: G>U>A>C (UGA), G>A>U>C (UAA) и A>U>C>G (UAG). Третий от стоп-кодона нижерасположенный нуклеотид оказывает лишь слабое влияние на эффективность терминации. Эти и другие такого рода данные были интерпретированы в пользу участия следующего за стоп-кодом нуклеотида во взаимодействии eRF-факторов с терминирующей рибосомой.

Результаты исследования возможной роли нуклеотидов, предшествующих терминирующему кодону, в терминации трансляции интерпретировать труднее, так как соответствующие замены могут непосредственно изменять первичную структуру С-концевой последовательности строящихся полипептидов. Такие эффекты были

обнаружены у *E. coli*: по крайней мере, две последние аминокислоты оказывают влияние на эффективность терминации синтеза соответствующего белка. У этой бактерии перед кодоном UGA чаще встречаются кодоны UCC(Ser), а также UUC и UUU(Phe), в то время как перед кодоном UUA – кодон AAG(Lys). Эти особенности объясняют возможным участием пептидил-тРНК, находящейся в Р-участке рибосом, или самой С-концевой аминокислоты в функционировании RF-факторов в А-участке. Зависимость терминации от природы аминокислот в С-концевой части строящегося полипептида была прямо продемонстрирована для *E. coli* и дрожжей.

2.5.5. Трансляция в митохондриях

Митохондрии являются органеллами эукариотических клеток, в которых в результате окислительного фосфорилирования энергия химических связей, освобождающаяся при метаболизме, накапливается в виде энергии макроэргических связей АТФ. Эти органеллы обладают собственным геномом – двухцепочечной ковалентно замкнутой митохондриальной ДНК (мтДНК), которая присутствует в каждой митохондрии в виде нескольких идентичных копий. В настоящее время наибольшее распространение получила гипотеза об эндосимбиотическом происхождении митохондрий, в соответствии с которой современные митохондрии животных берут свое начало от α -протеобактерий (к которым принадлежит современная *Rickettsia prowazekii*), внедрившихся в цитозоль клеток-предшественников. Считается, что за время эндосимбиоза бактерии передали большую часть своих жизненно важных генов хромосомам клетки-хозяина, сохранив в своем геноме (в случае клеток человека) информацию лишь о 13 полипептидах, 22 тРНК и двух рРНК. Все полипептиды входят в состав ферментативных комплексов системы окислительного фосфорилирования митохондрий.

Как уже можно видеть из размера мтДНК (ее длина у человека составляет 16569 п.о.), митохондрии не являются самовоспроизводящимися генетическими системами. Репликация и транскрипция их генома зависят от *транс*-действующих факторов, кодируемых ядерным геномом. Все митохондриальные тРНК акцептируют аминокислотные остатки при участии аминоацил-тРНК-синтетаз, импортируемых в митохондрии из цитоплазмы. В цитоплазме же клеток позвоночных синтезируются все рибосомные белки

митохондрий, мРНК которых транскрибируются с ядерных генов. Даже происхождение белков комплексов системы окислительного фосфорилирования является смешанным – ядерно–митохондриальным. Все полипептиды, кодируемые ядерным геномом, синтезируются рибосомами цитозоля. Как правило, они обладают N-концевой отщепляемой сигнальной последовательностью, которая обеспечивает направленный перенос в митохондрии.

Декодирование генетической информации в митохондриях.

Сравнение первичной структуры мтДНК с последовательностями аминокислот митохондриальных белков выявило ряд особенностей генетического кода, используемого митохондриями. Имеются вариации и в частоте использования кодонов митохондриями разных видов. В мтДНК большинства филогенетических групп триплет TGA кодирует Trp, хотя обычно на нем происходит терминация синтеза полипептидных цепей. С другой стороны, кодон AGR (где R=A или G) в мтДНК позвоночных является терминирующим, у иглокожих кодирует Ser, а у дрожжей, как обычно, – Arg.

Другой характерной особенностью генетической системы митохондрий является упрощение самого механизма декодирования информации, заключенной в митохондриальной мРНК. Интересно, что трансляция всех кодонов в митохондриях происходит с участием далеко не всех 32 видов молекул тРНК, которые обычно требуются для осуществления биосинтеза белка в соответствии с гипотезой неоднозначного соответствия Крика. Например, лишь 22 видов молекул тРНК достаточно для трансляции всех мРНК 13 генов мтДНК человека. Это становится возможным благодаря использованию основания U в первом (wobble) положении антикодона тРНК одного вида для распознавания всех четырех кодонов каждого из семейств, кодирующих конкретную аминокислоту.

Миторибосомы. Одним из первых указаний на прокариотическое происхождение митохондрий было отличие набора антибиотиков, подавляющих трансляцию в этих органеллах, от антибиотиков, которые ингибируют биосинтез белка в цитозоле клеток эукариот. Дальнейшие исследования показали, что некоторые компоненты системы трансляции митохондрий гомологичны соответствующим компонентам бактериальных систем.

Рибосомы митохондрий, или *миторибосомы*, ассоциированы с митохондриальным матриксом. На основании определения стационарного уровня митохондриальных рРНК в гепатоцитах крыс было установлено, что отдельная митохондрия содержит <100 миторибосом. Структурные исследования миторибосом выявили их существенные отличия от цитоплазматических и бактериальных рибосом. Для миторибосом животных характерно очень низкое содержание РНК и, как следствие, более низкий коэффициент седиментации (~55S). Коэффициенты седиментации большой и малой субчастиц миторибосом составляют соответственно ~39S и ~28S, которые содержат в своем составе 16S и 12S рРНК. Небольшая 5S рРНК, характерная для обычных рибосом, в миторибосомах отсутствует. Тем не менее, участок 3'-конца 16S рРНК миторибосом человека длиной в 23 нуклеотида обнаруживает 68% гомологию с 5S рРНК *Bacillus subtilis*. Полагают, что эта часть 16S рРНК представляет собой укороченный функциональный аналог 5S рРНК бактериальных рибосом.

Низкая массовая доля РНК в миторибосомах компенсируется повышенным содержанием белков. В результате суммарная молекулярная масса миторибосом приближается к таковой рибосом бактерий. Полагают, что часть белков миторибосом могла взять на себя функции недостающих последовательностей рРНК. С помощью двумерного электрофореза удастся обнаруживать 85 (быки) или 86 (крысы) миторибосомных белков. Тем не менее, истинное количество белков, составляющих миторибосомы, остается неизвестным, поскольку часть электрофоретически выявляемых полипептидов может представлять собой продукты протеолитической деградации или быть предшественниками их зрелых форм.

Биосинтез белка в митохондриях. Несмотря на то что изолированные митохондрии активно синтезируют белок, до сих пор из них не удается получить эффективные бесклеточные системы трансляции. Это сдерживает исследование молекулярных механизмов трансляции в митохондриях, и имеется мало информации относительно факторов трансляции, участвующих в синтезе митохондриальных белков.

Трансляция в митохондриях уникальна во многих отношениях. Как уже упоминалось, одной из особенностей биосинтеза белка в митохондриях является малое число видов рРНК и тРНК, вовлеченных в этот процесс.

Остается непонятным, каким образом миторибосомы распознают иницирующие кодоны. Поскольку митохондриальные мРНК лишены 5'UTR, которые характерны для бактериальных и ядерных мРНК, взаимодействие миторибосом с иницирующими кодонами должно определяться другими принципами, отличающимися от сканирования цитоплазматическими рибосомами эукариот или от взаимодействий TIR-последовательностей мРНК с 16S рРНК у бактерий. Обнаруженная низкая эффективность трансляции митохондриальных мРНК *in vitro* может быть следствием строения их 5'-концевых последовательностей, и для осуществления биосинтеза митохондриального белка на приемлемом уровне требуется присутствие мРНК в высоких концентрациях, что и имеет место в нативных митохондриях.

В бесклеточных системах было показано, что 28S субчастица миторибосом обладает способностью прочно связываться с мРНК независимо от последовательности нуклеотидов в отсутствие дополнительных факторов и инициаторной тРНК, что отличает ее от субчастиц других про- и эукариотических рибосом. Для ее эффективного взаимодействия с мРНК длина полинуклеотида должна быть ~400 нуклеотидов, тогда как сама субчастица защищает от действия РНКазы T_1 участок мРНК длиной в 30–80 нуклеотидов. Возможно, это объясняет, почему самые короткие ОРС мтДНК являются перекрывающимися, так как благодаря такому строению соответствующие мРНК приобретают размеры, необходимые для оптимального взаимодействия с малой субчастицей миторибосом. Полагают, что после взаимодействия с мРНК малая субчастица при участии еще не идентифицированных факторов трансляции перемещается к иницирующему кодону на 5'-конце матрицы.

В настоящее время единственным охарактеризованным митохондриальным фактором инициации трансляции является mtIF2, который обеспечивает GTP-зависимое вхождение инициаторной формилметионил-тРНК (fMet-тРНК) в комплекс, образованный малой субчастицей миторибосом и мРНК. Вероятно, гидролиз GTP сопровождается освобождением mtIF2 из тройного комплекса с последующим соединением малой и большой (39S) субчастиц миторибосом с образованием комплекса, способного к элонгации полипептидных цепей.

Из печени быка были выделены три фактора элонгации трансляции: mtEF-Tu, mtEF-Ts и mtEF-G, которые обладают высокой гомологией с

соответствующими прокариотическими факторами. На этом основании полагают, что механизм элонгации строящихся полипептидных цепей в митохондриях в основных чертах соответствует таковому *E. coli*, уже подробно рассмотренному выше. В отличие от бактериальных факторов EF-Tu и EF-Ts соответствующие факторы митохондрий образуют друг с другом более прочный комплекс, для диссоциации которого недостаточно присутствия только GTP, однако комплекс диссоциирует при наличии в бесклеточных системах GTP и аминоацилированной тРНК.

Патологические последствия мутационных повреждений системы трансляции митохондрий. Точковые мутации, обнаруживаемые в мтДНК, как правило, являются *гетероплазматическими*, т.е. в одной и той же клетке присутствуют как нормальные, так и мутантные мтДНК. Именно это явление позволяет существовать клеткам с мтДНК, содержащими летальные мутации, которые в обычных условиях были бы элиминированы. Энцефаломиопатии, для которых характерен материнский тип наследования (с цитоплазмой яйцеклетки), часто вызываются точковыми мутациями в генах митохондриальных тРНК. В частности, такими патологическими последствиями сопровождаются транзиции A3243G и A8344G в генах тРНК^{Leu(UUR)} и тРНК^{Lys} соответственно. Предполагается, что в первом случае нарушается правильный процессинг первичного полицистронного транскрипта, что приводит к образованию гибридной молекулы с последовательностями 16S рРНК–тРНК^{Leu(UUR)}–ген *ND1*, которая, из-за присутствия в ней последовательности 16S рРНК, включается в нефункциональную субчастицу рибосом. Альтернативно, обсуждаемые мутации могут дестабилизировать молекулы тРНК.

2.5.6. Трансляция в хлоропластах.

Хлоропласты являются органеллами клеток растений, осуществляющих процесс фотосинтеза – преобразование энергии квантов света в энергию макроэргических связей АТФ. Так же как и митохондрии, хлоропласты обладают собственным геномом, представленным множественными копиями кольцевой ковалентно замкнутой ДНК (хпДНК – *ctDNA*), длина которой обычно составляет ~150 т.п.о. Геном хлоропластов включает в себе более 100 различных генов. В соответствии с теорией эндосимбиоза хлоропласты произошли от

цианобактерии *Anacystis nidulans* (*Synechococcus* PCC6301), которая в ходе адаптации к внутриклеточному существованию передала основную часть своих генов хромосомам ядра клетки-хозяина. В результате образовавшийся хлоропласт стал зависимым от ядра в отношении биосинтеза импортируемых хлоропластных белков и генетического контроля экспрессии собственных генов. Как и митохондрии, хлоропласты обладают собственной системой транскрипции и трансляции, а также репликации хлДНК.

В отличие от митохондрий животных, система трансляции хлоропластов высоко гомологична системе бактерий и представлена 70S рибосомами, собственными тРНК и аминоацил-тРНК-синтетазами, многочисленными факторами трансляции и т.п. Геном хлоропластов содержит гены всех рРНК (16S, 23S и 5S), которые кластеризованы и транскрибируются полицистронно. В большой субчастице рибосом хлоропластов рРНК 23S-типа часто представлена двумя или четырьмя фрагментами. Так, в рибосомах *Chlamydomonas eugametos* она представлена четырьмя фрагментами длиной 280 (α-фрагмент), 52 (β), 810 (γ) и 1720 (δ) нуклеотидов. Вторичная структура этих фрагментов практически идентична предсказанной структуре соответствующих участков 23S рРНК *E. coli*. На этом основании делается вывод, что физическая непрерывность молекулы 23S рРНК не существенна для ее функционирования.

70S рибосомы хлоропластов содержат ~60 рибосомных белков, что превышает их содержание (55 полипептидов) в рибосомах *E. coli*. Приблизительно 1/3 рибосомных белков кодируется хлДНК, а 2/3 – ядерным геномом. Рибосомы хлоропластов высших растений содержат, по крайней мере, пять белков, не имеющих гомологов в рибосомах *E. coli*.

Геномы хлоропластов, для которых определена полная первичная структура, содержат 27–35 потенциальных генов тРНК. При этом в геноме для кодирования полипептидов используются все теоретически возможные кодоны (61). Это приводит к ситуации, характерной и для митохондрий, – у хлоропластов отсутствует полный набор тРНК, необходимых для декодирования этих кодонов. В данном случае проблема, по-видимому, решается так же, как и у митохондрий: индивидуальные тРНК^{Pro}(UGG), тРНК^{Ala}(UGC) и тРНК^{Arg}(ACG) распознают по четыре кодона, которые кодируют каждую из аминокислот, акцептируемых соответствующими молекулами тРНК (в скобках представлены последовательности антикодонов тРНК).

Генетическая информация хлДНК во многих случаях редактируется на уровне мРНК (подробно о механизме см. раздел 2.2.2). В этом случае в результате запрограммированных замен нуклеотидов в мРНК происходит создание новых иницирующих и терминирующих кодонов, а также изменение их смысла. Лишь очень редко редактирование сопровождается синонимическими заменами нуклеотидов (без изменения смысла кодона). Редактирование является критическим событием в экспрессии генов хлоропластов, так как неотредактированные транскрипты не способны правильно транслироваться. Очень эффективным результатом редактирования предшественника мРНК хлоропластов является создание в результате двух замен нуклеотидов иницирующего и терминирующего кодонов с образованием новой открытой рамки считывания, т.е. фактически нового гена, посттранскрипционно. В хлоропластах кукурузы, табака и черной сосны, геномы которых полностью секвенированы, имеется соответственно 26, 32 и 26 сайтов редактирования.

Зрелые и функционально активные мРНК хлоропластов не обладают кэп-группами и не полиаденилированы на 3'-концах. Из 70 генов, кодирующих белки в хлоропластах табака, лишь пять транскрибируются моноцистронно. Полицистронные предшественники мРНК подвергаются эндонуклеазному процессингу с образованием моноцистронных матриц. Более того, искусственные дицистронные мРНК не транслируются в бесклеточных системах. На этом основании делается вывод, что в хлоропластах в синтезе белка участвуют моноцистронные мРНК.

Среди 79 исследованных генов, кодирующих белки в хлоропластах табака, 30 содержат SD-подобные последовательности в 20-нуклеотидном участке перед иницирующим кодоном. Остальные 49 транскриптов также содержат такие последовательности, но их положение не фиксировано на матрице. Мутационные изменения некоторых SD-подобных последовательностей снижают эффективность трансляции мутантных мРНК, что указывает на функциональную значимость этих участков мРНК. Детальное исследование 5'UTR мРНК гена *psbA* хлоропластов табака позволило идентифицировать *цис*-действующие регуляторные элементы, существенные для ее трансляции. Два из них – RBS1 (AAG) и RBS2 (UGAU), расположенные между нуклеотидами в положениях –11 и –9, –25 и –22 соответственно

комплементарны 3'-концу 16S рРНК хлоропластов. Полагают, что они участвуют во взаимодействии 30S субчастицы рибосом с мРНК. AU-богатая последовательность, расположенная между ними (UAAAUAAA) и получившая название AU-бокса, также критична для трансляции. Возможно, с этой последовательностью взаимодействуют *транс*-действующие белковые факторы. На наличие таких факторов трансляции указывают многочисленные данные мутационного анализа *Chlamydomonas*.

Как и у бактерий, AUG является основным иницирующим кодоном мРНК хлоропластов и направляет включение в полипептидную цепь формилметионина. Информация о молекулярных механизмах отдельных этапов трансляции в хлоропластах еще не получена.

ГЛАВА 3. ОСНОВНЫЕ ПУТИ РЕГУЛЯЦИИ ЭКСПРЕССИИ ГЕНОВ

В этой главе будет рассмотрена лишь небольшая часть той почти необъятной информации, которая накоплена к настоящему времени в отношении механизмов регуляции экспрессии генов.

Конечным результатом экспрессии любого известного гена на молекулярном уровне является образование молекул РНК или белка, информация о первичной структуре которых закодирована в этом гене. Процесс биосинтеза белка складывается из многих взаимосвязанных этапов. Как уже упоминалось выше, основными из них являются транскрипция, трансляция, а также посттранскрипционные и посттрансляционные процессинг и модификации РНК и белка. Поэтому изменение скорости протекания каждого из данных этапов сопровождается, в конечном счете, изменением внутриклеточного содержания функционально активного продукта экспрессии гена. Следовательно, регуляторное воздействие на любом из этих этапов может привести к изменению уровня экспрессии соответствующего гена в клетках. Регулируемая экспрессия генов предполагает высокоспецифическое изменение внутриклеточного содержания кодируемых этими генами белков и нуклеиновых кислот в ответ на действие продуктов экспрессии других генов или регуляторных сигналов внутри- и внеклеточного происхождения, например низкомолекулярных метаболитов, ксенобиотиков или физических факторов (температура, ионизирующее излучение и т.п.). Избирательность таких воздействий становится возможной благодаря образованию высокоспецифических белок–белковых комплексов, комплексов лиганд–рецептор, распознаванию белками определенных последовательностей нуклеотидов ДНК или РНК, а также вследствие комплементарных взаимодействий нуклеиновых кислот друг с другом.

Избирательное действие низкомолекулярных биорегуляторов на гены происходит опосредованно через соответствующие рецепторы белковой природы. При этом, как правило, реализуется следующая схема: высоко- или низкомолекулярный эффектор (лиганд) специфически связывается с регуляторным белком-рецептором (например репрессором или активатором гена), изменяя конформацию рецептора таким образом, что он приобретает

способность распознавать регуляторные последовательности нуклеиновых кислот или других регуляторных белков. Подобные взаимодействия, происходящие на одном из вышеупомянутых этапов биосинтеза белка, далее сопровождаются изменением эффективности экспрессии его гена. Очевидно, что наиболее продуктивно можно оказывать влияние на экспрессию гена через его транскрипцию. При таком способе регуляции должен изменяться внутриклеточный уровень соответствующих мРНК, который может лимитировать биосинтез белков рибосомами. Кроме того, прекращение синтеза мРНК для уменьшения внутриклеточного содержания ненужных в данный момент белков экономично и с энергетической точки зрения, так как с прекращением транскрипции перестает затрачиваться энергия на биосинтез ненужных макромолекул – мРНК или их предшественников. Действительно, регуляция экспрессии генов на уровне транскрипции широко распространена в природе.

Однако этот способ не является единственным. В ряде случаев накопление мРНК в виде внутриклеточного пула без немедленной их трансляции происходит перед определенными стадиями дифференцировки клеток, например в яйцеклетках до оплодотворения. Эти неактивные мРНК могут длительное время храниться и немедленно использоваться после получения клетками соответствующих сигналов. Кроме того, альтернативный процессинг мРНК приводит к образованию из одного и того же предшественника нескольких зрелых мРНК, трансляция которых сопровождается синтезом разных белковых продуктов. Специальные регуляторные механизмы могут изменять соотношения таких процессированных мРНК и, как следствие, внутриклеточное содержание кодируемых ими полипептидов. Использование регуляции данного типа позволяет повысить кодирующие возможности генов путем более сжатого хранения генетической информации. Не меньшую пользу для клетки приносит и регуляция экспрессии генов на других уровнях: трансляционном и посттрансляционном, которые только в совокупности способны обеспечивать поддержание жизненно важного гомеостаза организма.

Это краткое резюме, в котором перечислены основные пути регуляции экспрессии генов на качественном уровне, подразумевает количественные изменения внутриклеточного содержания белков и РНК, закодированных в

геноме живого организма, в ответ на соответствующие регуляторные воздействия. Представляется весьма желательным построение количественных моделей регуляции активности генов. Такая постановка вопроса особенно актуальна с учетом тех глубоких перестроек фенотипа, которые сопровождают количественные изменения внутриклеточных концентраций продуктов конкретных генов (например в результате нарушения дозовой компенсации), в ряде случаев приводящие к развитию тяжелейших заболеваний (в частности синдрому Дауна).

3.1. Регуляция экспрессии генов на уровне транскрипции у прокариот

Регуляция транскрипции в клетках осуществляется на уровне индивидуальных генов, их блоков и даже целых хромосом. Возможность управления многими генами, как правило, обеспечивается наличием у них общих регуляторных последовательностей нуклеотидов, с которыми взаимодействуют однотипные факторы транскрипции. В ответ на действие специфических эффекторов такие факторы приобретают способность с высокой точностью связываться с регуляторными последовательностями генов. Следствием этого является ослабление или усиление транскрипции (репрессия или активация) соответствующих генов.

Три основных этапа транскрипции – инициация, элонгация и терминация, рассмотренные нами выше, используются бактериальными клетками для регуляции синтеза РНК. То же, по-видимому, характерно и для остальных живых организмов. Ниже приведены примеры регуляторных воздействий на транскрипцию.

3.1.1. Регуляция на уровне инициации транскрипции

Активность многих генов прокариот регулируется с помощью белковых факторов, взаимодействующих с регуляторными участками промоторов генов. При этом происходят как активация транскрипции генов, так и подавление считывания генетической информации РНК-полимеразами. В первом случае регуляторные белковые факторы называют активаторами, осуществляющими позитивную регуляцию транскрипции, а во втором – репрессорами. Регуляцию, связанную с подавлением транскрипции, называют негативной.

Механизмы, при помощи которых активаторы стимулируют инициацию транскрипции, могут быть рассмотрены с двух точек зрения – кинетической и структурной. Поскольку активация промоторов путем образования открытых комплексов является лимитирующей стадией на пути активации транскрипции в целом, действие различных активаторов может быть охарактеризовано по изменению (увеличению) значений кинетических параметров реакций, происходящих на разных этапах активации. Так, при действии активирующего комплекса Cgr-sAMP на lac -промотор происходит десятикратное увеличение равновесной константы ассоциации K_B РНК-полимеразы с промотором с образованием закрытого комплекса. Активация промотора P_{RM} фага λ , опосредованная cI -белком (см. ниже), характеризуется пяти–десятикратным увеличением константы скорости k_f перехода закрытых комплексов в открытые. Активирующее действие белка $\lambda\text{-cII}$ на промотор P_{RE} сопровождается изменением обоих вышеупомянутых кинетических параметров. Активация транскрипции может быть также опосредована увеличением скорости освобождения промотора РНК-полимеразой после инициации синтеза РНК.

Многие активаторы транскрипции, в том числе и Cgr-sAMP , сгибают молекулу ДНК после взаимодействия с ней, причем центр такого изгиба находится в сайте связывания активатора. Однако с использованием мутантных белков было установлено, что изгибание ДНК и связывание активаторов с ДНК как таковое еще не обеспечивают активацию транскрипции. В большинстве случаев абсолютно необходимым условием активации является наличие контакта между специфическими областями поверхностей молекул активатора и РНК-полимеразы, часто с ее α -субъединицами. Важным следствием образования контактов между активаторами и холоферментом РНК-полимеразы является часто наблюдаемый синергизм в связывании обоих белков с соответствующими промоторами. При этом мутации в сайтах связывания активаторов или промоторе как таковом могут предотвращать активацию транскрипции путем изменения конформации молекулы связанного активатора или контактного участка на РНК-полимеразе.

Последовательности нуклеотидов промоторных участков генов, с которыми взаимодействуют молекулы репрессора, получили название операторов. Во многих случаях репрессор связывается с оператором только в

присутствии низкомолекулярного лиганда, специфически взаимодействующего с репрессором. Такие низкомолекулярные эффекторы получили название корепрессоров. Они часто требуются и для функционирования белков-активаторов транскрипции. Простейший механизм репрессии заключается в стерическом блокировании связывания РНК-полимеразы с промотором. Это происходит в том случае, если последовательности нуклеотидов мест посадки РНК-полимеразы на промотор и репрессора на оператор перекрываются.

Некоторые бактериальные белки-репрессоры могут оказывать свое негативное действие на этапы инициации, происходящие после связывания РНК-полимеразы с промотором. Например, молекулы репрессора *gal*-оперона *E. coli*, связавшиеся с операторами O_E и O_I , центры последовательностей которых расположены соответственно на расстояниях $-60,5$ и $+53,5$ по отношению к точке инициации транскрипции, вызывают образование петли участка ДНК, заключенного между ними, но не препятствуют взаимодействию РНК-полимеразы с промотором. Они оказывают свое действие на последующие этапы инициации, предшествующие образованию первой фосфодиэфирной связи. В том случае, если лишь одна молекула репрессора связывается с внешним оператором O_E , он частично ингибирует транскрипцию путем взаимодействия с α -субъединицей РНК-полимеразы. Это сопровождается понижением уровня, но не полным прекращением синтеза РНК *gal*-оперона, т.е. более тонким изменением уровня экспрессии соответствующих генов.

Распространенным механизмом активации транскрипции с помощью белков-активаторов является облегчение ее инициации РНК-полимеразами после образования контакта между ферментом и белком-активатором, связанными с регуляторной областью промотора, что сопровождается конформационными изменениями РНК-полимеразы. У бактерий имеются белки-регуляторы, обладающие активностью как репрессора, так и активатора транскрипции. Такими "амфотерными" свойствами обладает, в частности, репрессор *ci* фага λ . Белок-активатор катаболических оперонов (*Crp*-белок) активирует транскрипцию бактериальных генов, продукты которых участвуют в расщеплении (катаболизме) различных органических соединений (преимущественно сахаров), используемых растущей бактериальной клеткой в качестве источника углерода. Свои свойства активатора *Crp*-белок приобретает лишь в комплексе с циклическим АМР (сАМР). Внутриклеточная концентрация

cAMP возрастает у бактерий, растущих на бедных питательных средах, и понижается в условиях избытка легко усвояемых источников углерода, например глюкозы. Поэтому система Csr–cAMP обеспечивает включение экспрессии катаболитных оперонов лишь на бедных питательных средах. Csr-белок может выступать и в роли репрессора транскрипции генов галактозного оперона *E. coli*. Если все гены катаболитных оперонов активируются Csr-белком в присутствии cAMP, то негативная регуляция их транскрипции происходит индивидуально. Хорошо известными примерами такого рода являются регуляции транскрипции *lac*-оперона *E. coli* под действием Lac-репрессора, а также галактозного и арабинозного оперонов специфическими белками-репрессорами этих оперонов.

Низкомолекулярные эффекторы могут изменять активность РНК-полимеразы не только опосредованно через белки-регуляторы, но и непосредственно при взаимодействии с ферментом. С помощью гуанозинтетрафосфата (ppGpp) в клетках *E. coli* осуществляется координация экспрессии генов рибосомных РНК (рРНК) и белков. Этот необычный нуклеотид (известный также под названием "магического пятна") синтезируется бактериальными клетками в условиях внутриклеточного недостатка аминокислот, что приводит к значительному снижению интенсивности транскрипции генов рРНК и белков и одновременной стимуляции синтеза РНК оперонов, контролирующих биосинтез аминокислот. В присутствии ppGpp очищенная РНК-полимераза прекращает синтез рРНК с одного из двух промоторов этих оперонов, что приводит к ослаблению, но не полному прекращению их транскрипции. Известны мутации, локализованные в гене β -субъединицы РНК-полимеразы *E. coli*, приводящие к прекращению влияния ppGpp на синтез РНК, что подтверждает наличие непосредственного контакта между нуклеотидом и ферментом в процессе транскрипции.

Некоторые регуляторные элементы бактерий, участвующие в активации транскрипции, так же как и энхансеры эукариот (см. раздел 3.2.2), могут располагаться на большом расстоянии (нескольких сотен нуклеотидов) от промоторов, на которые они оказывают свое действие. В этом случае контакт активатора с РНК-полимеразой обеспечивается благодаря выпетливанию участка ДНК, расположенного между данными регуляторными элементами, что приводит к пространственному сближению двух белков. Прямое доказательство

образования таких петель на ДНК *E. coli* было, в частности получено с помощью электронной микроскопии для белков NtrC и NifA, действующих соответственно на промоторы генов *glnA* и *nifH*. Другим путем достижения белком-активатором молекулы РНК-полимеразы на удаленном промоторе (например промоторе поздних генов бактериофага T4) является его перемещение вдоль отрезка ДНК, разделяющего эти два регуляторных элемента. Процесс такого перемещения может быть инициирован последовательностями нуклеотидов, расположенными выше или ниже промотора на расстоянии нескольких сотен пар оснований.

Одним из давно обсуждающихся вопросов является необходимость изменения структуры ДНК в окрестностях промоторов под действием белков-активаторов для активации транскрипции. В ряде случаев такие доказательства были получены. Так, в *mer*-локусе *E. coli*, обеспечивающем устойчивость бактериальных клеток к ионам ртути, связывание Мер-белка с регуляторным участком промотора (*merT*) в присутствии ртути сопровождается раскручиванием спирали ДНК в районе промотора на $\sim 50^\circ$. Это приводит к образованию правильного расстояния между сайтами связывания активатора и промотором, так как первый расположен необычно – между нуклеотидами в положениях -35 и -10 промотора. Без такого изменения структуры ДНК связавшийся с ним активатор не может образовать правильного контакта с РНК-полимеразой.

В заключение этого раздела следует еще раз подчеркнуть, что у активируемых бактериальных промоторов образование открытых комплексов в отсутствие активаторов является лимитирующей стадией при инициации транскрипции. Первичная структура активируемых промоторов весьма слабо соответствует каноническим структурам. При этом мутации, обеспечивающие функционирование таких промоторов без активации, увеличивают скорость образования открытых комплексов РНК-полимеразой.

3.1.2. Регуляция синтеза РНК на уровне элонгации и терминации

Выше было отмечено, что РНК-полимераза в процессе элонгации цепей РНК перемещается неравномерно вдоль матричной ДНК и во время ее движения имеют место остановки (паузы). Время задержки молекул РНК-

полимеразы в определенных участках транскрибируемых генов меняется под действием различных белковых факторов. При этом эффективность транскрипции соответствующих фрагментов ДНК зависит от последовательностей нуклеотидов, окружающих транскрибируемые участки генов.

Основная регуляторная роль терминаторов транскрипции заключается в прекращении синтеза РНК на границе гена и освобождении полученной РНК из транскрипционного комплекса. Механизмы функционирования терминаторов уже были кратко рассмотрены в разделе, посвященном транскрипции. Однако терминаторы встречаются не только на границах одиночных генов, но и в конце генов, входящих в состав оперонов. Эффективность терминации транскрипции на таких внутренних терминаторах может регулироваться, что сопровождается изменением скорости синтеза РНК на последовательностях нуклеотидов оперонов, расположенных за терминаторами.

Функционирование аттенуаторов – регулируемых терминаторов транскрипции бактерий, которые были рассмотрены в разделе 2.1.3, сопряжено с синтезом лидерного пептида рибосомами. Этот тип регуляции используется грамотрицательными бактериями для изменения уровня транскрипции многих оперонов, контролирующих биосинтез аминокислот. Как уже упоминалось, отличительной чертой такого типа регуляции является образование альтернативных вторичных структур РНК, которые формируются под влиянием рибосом, прекращающих трансляцию на кодонах аминокислот, которые клеткам необходимо синтезировать. В дополнение к этому, как прокариоты, так и эукариоты обладают способностью реагировать на многие внеклеточные и внутриклеточные процессы изменением скорости элонгации транскриптов.

Влияние условий роста бактериальных клеток на элонгацию цепей РНК. Механизмы преждевременной терминации транскрипции и антитерминации интенсивно используются бактериями для регуляции внутриклеточного метаболизма при изменении условий роста клеток. В частности, скорости синтеза рРНК и тРНК строго координированы со скоростью роста, и наоборот, число делений клеток в единицу времени прямо зависит от скорости транскрипции соответствующих оперонов. Совокупность событий, развивающихся у бактерий в ответ на изменение скорости их роста, например при изменении содержания питательных веществ в среде, получила название

строного ответа (*stringent response*). Первичной сигнальной молекулой строгого ответа является гуанозинтетрафосфат – ppGpp, который уменьшает скорость элонгации РНК-полимеразой и вызывает преждевременную терминацию транскрипции в *glnB*-опероне.

Транскрипция оперонов рибосомных белков также контролируется на уровне элонгации. В частности, у *E. coli* транскрипция оперона, кодирующего рибосомный белок S10 и десять других рибосомных белков, контролируется через задержку транскрипции, для реализации которой используется продукт третьего гена оперона – рибосомный белок L4. Этот белок взаимодействует со шпилькой синтезируемой РНК, вызывая преждевременную терминацию транскрипции в лидерном участке гена S10 – первом гене оперона. В отличие от классической аттенюации альтернативные вторичные структуры РНК в этом случае не являются причиной изменения эффективности транскрипции соответствующего участка ДНК. Белок L4 связывается со шпилькой на 70–80 нуклеотидов выше сайта терминации транскрипции. Шпилька в синтезируемой РНК образуется выше сайта терминации независимо от присутствия белка L4, однако без него молекула РНК-полимеразы не замечает сигнала терминации транскрипции. Этот механизм предотвращает накопление в клетке свободных рибосомных белков, не включившихся в состав рибосомных субчастиц.

Транскрипция оперона биосинтеза пиримидинов у *E. coli* регулируется как на уровне элонгации, так и на этапе освобождения промотора. Первый тип регуляции является еще одним примером классической аттенюации, тогда как во втором случае реализуется нетривиальный механизм. Оперон индуцируется в присутствии низких внутриклеточных концентраций UTP. При повышении ее содержания транскрипционный комплекс продолжает эффективно формироваться на промоторе, однако во время инициации РНК-полимераза начинает реитеративный синтез гомополимера. Синтезируются длинные цепи поли(U), а в продуктивную фазу синтеза РНК фермент не вступает, потому что не может покинуть промотор. Недавно тот же механизм контроля был описан для оперона *codBA* *E. coli*. По мнению М. Чамберлина (1997 г.) аналогичный механизм может функционировать и у эукариот.

В отличие от только что рассмотренного механизма контроля пиримидинового оперона регуляция транскрипции оперона *purB* *E. coli* осуществляется на уровне элонгации РНК с помощью пуринового репрессора –

ДНК-связывающего белка, взаимодействующего со специфической последовательностью нуклеотидов. Оперон *purB* является одним из девяти *pur*-оперонов, и два других гена также контролируются пуриновым репрессором. Однако только в опероне *purB* репрессор взаимодействует с ДНК значительно ниже сайта инициации транскрипции и создает препятствие элонгирующей РНК-полимеразе. Этот эффект не зависит от *purB*-промотора и не сопряжен с трансляцией.

У *B. subtilis* регуляция транскрипции оперонов биосинтеза пуринов и пиримидинов также связана с задержкой синтеза РНК, однако используемый при этом механизм отличается от такового у *E. coli*. В этом случае первичными регуляторами транскрипции являются уже не рибосомы, а новые белки, которые распознают другие элементы вторичной структуры РНК. В частности, было предсказано образование альтернативных шпилек в РНК пуринового оперона под действием РНК-связывающего белка, ассоциированного с гуанином. Реализуемый механизм оказался аналогичным механизму репрессии биосинтеза триптофана, о котором подробнее речь пойдет ниже.

Из общих соображений можно предположить, что биосинтез предшественников РНК, необходимых для синтеза пуринов и пиримидинов, должен оказывать прямое влияние на транскрипцию через изменение уровней соответствующих нуклеотидов. Действительно, у дрожжей, дефектных по фактору элонгации EFIS, стимулирующему элонгацию РНК у эукариот, наблюдается повышенная чувствительность к 6-азаурацилу, который уменьшает внутриклеточный пул GTP и UTP. К такому же фенотипу приводили и некоторые мутационные изменения субъединиц РНК-полимеразы II *S. cerevisiae*. Для некоторых, хотя и не для всех мутантных РНК-полимераз этого типа, были характерны нарушения в элонгации транскриптов или взаимодействии с факторами элонгации *in vitro*. Все это еще раз указывает на универсальный характер тонкой координации в клетках метаболизма нуклеотидов и синтеза РНК.

Регуляция транскрипции не всех бактериальных оперонов биосинтеза аминокислот у бактерий следует сценарию аттенуации, характерному для *trp*-, *his*-, *leu*- и *ilv*-оперонов *E. coli*. В частности, транскрипция *trp*-оперона *B. subtilis* также регулируется на уровне элонгации, однако в этом случае в регуляции участвует РНК-связывающий белок TRAP, который, кроме того,

взаимодействует с триптофаном. После образования комплекса с триптофаном белок TRAP приобретает способность связываться с РНК оперона в ее лидерном участке. Такое взаимодействие модифицирует вторичную структуру РНК, что способствует терминации транскрипции. Более того, белок TRAP может оказывать негативное влияние на трансляцию полицистронной мРНК оперона при повышении внутриклеточного содержания триптофана. В отличие от *E. coli* на процесс транскрипции *trp*-оперона у *B. subtilis* рибосомы не оказывают прямого влияния.

Сопряжение транскрипции и репарации ДНК. Многие повреждения ДНК вызывают прекращение элонгации транскриптов у бактерий и эукариот, вызывая переход элонгирующего транскрипционного комплекса в неактивное состояние. При этом преимущественно репарируются транскрибируемые цепи ДНК. Обнаружены белковые факторы, осуществляющие сопряжение транскрипции с репаративным синтезом ДНК. В частности, бактериальный белковый комплекс UvrAB, участвующий в репаративной эксцизии (вырезании) нуклеотидов (NER – см. раздел 5.2.2), может непосредственно взаимодействовать с β -субъединицей РНК-полимеразы, а также с одноцепочечными участками ДНК в составе открытых промоторных и тройных элонгирующих комплексов. У *E. coli* некоторые из этих воздействий обеспечивают контакты между молекулами РНК-полимеразы, прекратившей элонгацию цепей РНК, и репарирующим комплексом, который с помощью фактора, сопрягающего транскрипцию и репарацию, выводит элонгирующий комплекс из состояния прекращения транскрипции. Не все повреждения ДНК оказывают влияние на элонгацию транскриптов РНК-полимеразами. Например, РНК-полимеразы как фага SP6, так и *E. coli* эффективно преодолевают в ДНК бреши, не содержащие азотистых оснований. При прохождении таких участков ДНК молекулы РНК-полимераз обычно включают остатки AMP в элонгируемую РНК независимо от матрицы.

Контроль элонгации РНК у бактериофагов. Во время вирусной инфекции размножающиеся вирусы овладевают контролем над экспрессией генов клетки-хозяина и используют ее для своих собственных нужд, что является общим свойством всех внутриклеточных паразитов и симбионтов. Многие из бактериофагов осуществляют контроль транскрипции на уровне элонгации.

Бактериофаг λ . Механизмы контроля транскрипции на уровне элонгации РНК у этого бактериофага изучены лучше, чем у других вирусов. Ключевые регуляторные фаговые белки N и Q контролируют транскрипцию всего фагового генома, обеспечивая антитерминацию транскрипции в местах обычного прекращения синтеза вирусных РНК, т.е. на терминаторах транскрипции. В результате происходит транскрипция всех генов, необходимых для размножения бактериофага, и он вступает на вирулентный путь развития, приводящий к лизису бактериальных клеток и выходу зрелых фаговых частиц в окружающую среду. Механизм антитерминации, обеспечиваемой N-белком, отличается от механизма Q-зависимой антитерминации. N-Белок осуществляет свои функции при прохождении элонгирующей РНК-полимеразой особых последовательностей ДНК (бокс А и бокс В), обеспечивающих сборку комплекса элонгирующей РНК-полимеразы и N-белка. Белок N сам по себе является РНК-связывающим белком, и его взаимодействие с РНК стабилизируется Nus-белками *E. coli*. Его объединение с РНК и комплексом сопутствующих белков заставляет элонгирующую РНК-полимеразу не замечать сигналы терминации на ДНК, как зависимые, так и не зависящие от факторов терминации транскрипции.

Во время антитерминации транскрипции, опосредуемой Q-белком, последний взаимодействует с нетранскрибируемой цепью ДНК промоторного или элонгирующего комплексов в пределах первых 20–25 нуклеотидов ниже точки инициации транскрипции позднего фагового промотора. При этом белок Q не образует комплекса с РНК и оказывает свое действие на транскрипцию через регуляторную субъединицу РНК-полимеразы σ^{70} . Для реализации Q-зависимой антитерминации *in vivo* необходимо, чтобы произошла задержка в перемещении элонгирующей РНК-полимеразы вдоль матрицы вскоре после инициации транскрипции. В это время уже освободившаяся из транскрипционного комплекса σ -субъединица входит в контакт с нетранскрибируемой цепью ДНК в открытом участке (транскрипционном пузырьке), расположенном на 15 нуклеотидов ниже точки инициации транскрипции. Таким образом, σ -субъединица остается ассоциированной с транскрибирующим комплексом и опосредует антитерминирующее действие Q-белка на больших расстояниях (несколько тысяч пар оснований) от промотора.

В этом случае совместное действие фактора инициации транскрипции (σ -субъединицы) и ДНК-связывающего белка-антитерминатора изменяют свойства элонгирующей РНК-полимеразы, которая освобождает промотор после кратковременной паузы в элонгации. Эти необычные свойства σ -субъединицы, проявляемые во время вирусной инфекции, позволяют высказывать предположение о ее возможном участии в контроле транскрипции бактериальных генов на уровне освобождения промотора и элонгации синтеза РНК.

Лямбдоидный бактериофаг HK022. При заражении клеток *E. coli* этим бактериофагом происходит исключение (подавление развития) фага λ , в котором участвует вирусный белок *nup*. Этот белок, родственник белку N фага λ , также участвует в контроле элонгации РНК у бактериофага HK022. Во время λ -инфекции *Nup*-белок нарушает функционирование N-белка, предотвращая антитерминацию синтеза РНК и вызывая его терминацию. При этом оба белка взаимодействуют с одними и теми же последовательностями РНК, однако оказывают противоположное действие на элонгирующий транскрипционный комплекс. Бактериофаг HK022 обладает собственной регуляторной системой антитерминации транскрипции, для функционирования которой не требуется белок *Nup*, но необходима его мРНК. В этом случае элонгирующий комплекс перестает узнавать терминаторы транскрипции после прямого взаимодействия *Nup*-РНК с β' -субъединицей РНК-полимеразы. Для осуществления антитерминации по такому механизму не требуется участия других вирусных белков.

Бактериофаг T4. Бактериофаг T4 использует бактериальную РНК-полимеразу для транскрипции своего собственного генома, ингибируя синтез РНК бактериальной клетки-хозяина. В подавлении транскрипции участвует фаговый белок *Alc*, взаимодействующий с β -субъединицей РНК-полимеразы *E. coli*, что сопровождается прекращением транскрипции матричных ДНК, содержащих остатки цитозина. T4-ДНК содержит вместо остатков C остатки 5-гидроксиметилцитозина и эффективно транскрибируется РНК-полимеразой как в присутствии *Alc*-белка, так и без него. В отличие от этого матричные ДНК, содержащие немодифицированные остатки C, не транскрибируются бактериальной РНК-полимеразой только в присутствии белка *Alc*, и даже

простое метилирование остатков цитозина в положении С5 обеспечивает защиту транскрипции от преждевременной терминации, вызываемой этим белком.

Координация элонгации транскриптов с метаболизмом ДНК.

Транскрипция происходит координированно с репликацией бактериальной ДНК и сегрегацией хромосом в дочерние клетки. Транскрибирующая РНК-полимераза часто встречается в процессе синтеза РНК с репликативным бактериальным комплексом. Для сведения к минимуму отрицательного эффекта такой встречи активно транскрибируемые гены *E. coli*, других бактерий и бактериофагов ориентированы на хромосомах таким образом, чтобы синтез ДНК и РНК происходил в одном направлении. Но даже в этом случае встреча двух работающих комплексов неизбежна, поскольку репликация ДНК происходит в 10–20 раз быстрее транскрипции тех же ее участков. Теоретически отрицательные последствия такого столкновения заключаются в том, что в отличие от ДНК-полимераз, которые функционируют по дистрибутивному механизму, т.е. способны продолжать синтез ДНК после его временного прекращения и отделения фермента от матрицы, РНК-полимеразы являются процессивными ферментами. Они не могут продолжить синтез РНК после распада элонгирующего комплекса и для вовлечения в новый цикл транскрипции требуют инициации синтеза РНК на промоторе. При этом энергия, затраченная на синтез недостроенного фрагмента РНК, безвозвратно теряется. Данная проблема особенно актуальна для больших эукариотических генов, например гена дистрофина, размер которого превышает 2000 т.п.о. Тем не менее, методом электронной микроскопии показано, что у *E. coli* репликативная вилка может смещать с ДНК элонгирующий транскрипционный комплекс. Установлено также, что элонгация транскриптов оказывает влияние на правильную терминацию репликации.

У бактериофага Т4 репликативный аппарат может функционировать без разрушения элонгирующих транскрипционных комплексов независимо от направления транскрипции по отношению к направлению синтеза ДНК. В опытах Б. Лиу и Б.М. Албертса (1995 г.) с высокоочищенными компонентами аппаратов транскрипции и репликации бактериофага Т4 было установлено, что при одновременно происходящих транскрипции и репликации оба работающих комплекса не мешают друг другу выполнять свои функции. При прохождении

репликативной вилки через транскрибируемый участок ДНК этого вируса положение РНК-полимеразы на матрице не изменяется. Во время контакта работающей репликативной вилки с РНК-полимеразой в составе тройного комплекса не было обнаружено изменений его стабильности как во время активного синтеза РНК, так и в состоянии задержки транскрипции. Неожиданно оказалось, что при столкновении репликативного и транскрипционного комплексов друг с другом РНК-полимераза переключается на использование в качестве матрицы вновь синтезированной дочерней цепи ДНК. Во время такой смены матриц РНК-полимераза сохраняет связь с растущим транскриптом, остается в активном состоянии и продолжает безошибочно синтезировать правильную цепь РНК. Этот пример указывает на возможный механизм координации репликации и транскрипции у других, более сложноустроенных организмов.

Белковые факторы, регулирующие элонгацию РНК у бактерий.

Белковые факторы и небольшие молекулы, изменяющие свойства элонгирующих РНК-полимераз, столь же разнообразны, как и упомянутые выше механизмы регуляции транскрипции на уровне элонгации РНК. Это особенно относится к эукариотам, о которых речь пойдет в следующем разделе книги. Некоторые факторы стимулируют элонгацию через супрессию временной задержки РНК-полимераз на транскрибируемых матрицах или путем активации комплексов, полностью прекративших транскрипцию. Другие факторы обеспечивают терминацию или антитерминацию транскрипции. У эукариот имеются белки, изменяющие активность белковых компонентов элонгирующих комплексов, а также структуру транскрибируемого хроматина. Некоторые регуляторные белки осуществляют свое действие, связываясь непосредственно с растущим транскриптом или матричной ДНК, другие – через белок-белковые взаимодействия. Во многих случаях механизм действия регуляторных белков остается невыясненным.

У *E. coli* имеются уже упоминавшиеся выше белковые факторы GreA и GreB, которые выводят элонгирующий комплекс РНК-полимеразы из состояния полного прекращения синтеза РНК путем стимуляции эндонуклеазного отщепления 3'-концевой части элонгируемого транскрипта. Ген, кодирующий белок GreA, впервые был обнаружен генетическими методами как ген-супрессор температурно-чувствительной мутации в β' -субъединице РНК-

полимеразы *E. coli*. Инактивация генов *greA* или *greB* по отдельности с помощью мутаций не сопровождается выраженным изменением фенотипа мутантных бактерий, однако двойная мутация делает бактериальные клетки температурно-чувствительными. Для таких мутантных клеток была характерна быстрая реверсия к обычному нетемпературно-чувствительному фенотипу. На этом основании делается вывод о значительном преимуществе в росте бактериальных клеток, обладающих функциональными белками GreA и GreB, даже при перmissive температуре. Те же генетические свойства характерны и для фактора элонгации транскрипции TFIIS(SII) *S. cerevisiae* – функционального (но не структурного) гомолога вышеупомянутых бактериальных факторов.

Среди регуляторов элонгации транскрипции, действующих по типу терминции–антитерминции синтеза РНК, для полноты картины следует упомянуть уже обсуждавшиеся выше фактор терминции транскрипции ρ , а также белки-антитерминаторы транскрипции N и Q бактериофага λ .

3.2. Регуляция экспрессии генов на уровне транскрипции у эукариот

Несмотря на то что основные принципы регуляции транскрипции генов у прокариотических и эукариотических организмов остаются неизменными – через специфические взаимодействия белков и нуклеиновых кислот друг с другом, а также между собой, данный процесс у эукариот характеризуется рядом существенных особенностей. Это связано, прежде всего, с необходимостью поддержания координированной экспрессии эукариотических генов в более сложноорганизованной генетической системе. Достаточно вспомнить, что в организме человека гистологически различают, по крайней мере, 100 типов клеток, формирующих его органы и ткани. Для любого типа клеток характерен свой уникальный набор экспрессирующихся генов, которые начинают функционировать во время дифференцировки клеток-предшественников. Кроме того, сам процесс формирования органов и тканей сопровождается пролиферацией строго определенных групп клеток, а также упорядоченным во времени и пространстве перемещением клеток. Все эти особенности жизнедеятельности клеток высших организмов обеспечиваются функционированием их генов.

Гены высших организмов подразделяют по функциональному признаку на две большие группы: "гены домашнего хозяйства" (housekeeping genes) и "гены роскоши" (luxury genes). К первой группе относятся гены, функционирующие повсеместно, на всех стадиях жизненного цикла организма. Они обеспечивают процесс гликолиза, биосинтез аминокислот и нуклеотидов, катаболизм белков и т.п. Гены, относящиеся ко второй группе, экспрессируются лишь в специализированных клетках и являются маркерами дифференцированных состояний этих клеток.

Сложность жизненного цикла многоклеточных организмов накладывает свои требования на особенности функционирования их генов. В частности, большое число генов и даже целые блоки их функционируют лишь на определенных стадиях эмбриогенеза и не транскрибируются в клетках взрослого организма. У человека к ним относятся, например гены α -фетопротейна. Экспрессия этих генов в клетках взрослого организма свидетельствует о развитии патологического процесса, в частности злокачественных новообразований в печени. Еще более ярким примером такого рода является избирательная инактивация одной из X-хромосом у самок млекопитающих.

Тканеспецифический характер экспрессии генов роскоши обеспечивается различными механизмами. В этом случае ключевую роль играют специфические взаимодействия белковых факторов транскрипции с регуляторными последовательностями нуклеиновых кислот. Транскрипцию генов высших организмов осуществляют, по крайней мере, три различные РНК-полимеразы. При этом для промоторов каждой из них характерны специфические регуляторные последовательности нуклеотидов, с которыми взаимодействуют свои факторы транскрипции, изменяющие уровень транскрипции соответствующих генов (см. рис. 1.4).

В свою очередь, сами эукариотические факторы транскрипции реализуют новый, известный (в таком масштабе) только у эукариот механизм регуляции экспрессии генов на уровне транскрипции так называемого *комбинаторного* типа. Молекулы факторов транскрипции обладают консервативными доменами, которые дают им возможность осуществлять высокоспецифические белок-белковые и белково-нуклеиновые взаимодействия. В результате, *in vivo* происходит объединение факторов транскрипции и других регуляторных

белков, обладающих соответствующими доменами, в разных сочетаниях в большие регуляторные комплексы. Каждое новое сочетание факторов, число которых хотя и велико, но ограничено, придает комплексу уникальные регуляторные свойства, обеспечивая изменение специфичности его взаимодействия с регуляторными последовательностями ДНК и другими регуляторными белками аппарата транскрипции. В результате реализации такого механизма достигается беспрецедентная гибкость в модуляции уровней транскрипции эукариотических генов и соответственно в контроле экспрессии фенотипических признаков клетки и организма.

Не менее уникальна способность эукариот использовать для регуляции транскрипции своих генов изменения структуры хроматина. С помощью таких эффективных механизмов осуществляется репрессия и дерепрессия генов во время дифференцировки клеток, и соответствующее функциональное состояние отдельных генов, их больших массивов и даже целых хромосом может поддерживаться на протяжении всей жизни организма. Перестройки хроматина в окрестностях регуляторных участков генов происходят и в связи с более тонкой регуляцией их транскрипции.

Несмотря на то что изменение уровней транскрипции генов является одним из важнейших способов регуляции их экспрессии, такая стратегия – лишь одна из многих, используемых эукариотическими организмами для контроля биосинтеза, содержания и функционирования соответствующих продуктов генов: белков или нуклеиновых кислот. В процессе синтеза и после его завершения первичный транскрипт подвергается многочисленным посттранскрипционным модификациям и процессингу. Таким образом, генетической информации, заключенной в конкретном гене, недостаточно для полноценной экспрессии, и чтобы ген правильно функционировал, требуется координированная работа дополнительных генов, многие из которых активны не вблизи регулируемых генов, а в других тканях, удаленных от клеток-мишеней. Для осуществления такой передачи регуляторных сигналов на большие расстояния в организме присутствуют специальные системы, осуществляющие генерацию сигналов, перенос их к клеткам-мишеням, а также специфическое распознавание сигналов клетками, которым они адресованы.

3.2.1. Передача сигнала и вторичные мессенджеры

Жизнь любой клетки, включая глобальные процессы ее роста, деления и даже гибели, зависит от внешних регуляторных сигналов, которые она воспринимает. Такими сигналами могут быть физические воздействия (температура, ионизирующее и другое электромагнитное излучение) или многочисленные химические соединения. Хорошо изученными веществами, которые организм использует для регуляции жизнедеятельности клеток, являются, например стероидные гормоны, цитокины или факторы роста, которые, достигая клеток-мишеней, вызывают в них специфические метаболические изменения, связанные в том числе и с изменением экспрессии больших групп генов. Не менее сильный и часто также специфический ответ вызывают различные физиологически активные вещества экзогенного происхождения, например феромоны или токсины.

Все эти сигналы, передающиеся через соответствующие сигнальные молекулы, являются первичными по отношению к тем каскадам биохимических реакций, которые запускаются в клетках в ответ на их воздействие. Первичные сигналы распознаются клетками благодаря наличию у них специальных молекул-рецепторов белковой природы, взаимодействующих с первичными сигнальными молекулами или воздействиями физической природы. Первичный сигнал, как правило, не действует прямо на те метаболические процессы в клетке, для регуляции которых он предназначен. Вместо этого воспринимающий его рецептор инициирует образование в клетке промежуточных химических соединений, запускающих внутриклеточные процессы, воздействие на которые было целью первичного внеклеточного сигнала. Поскольку такие промежуточные соединения несут в себе информацию о первичном регуляторном сигнале и являются вторичными его переносчиками, они получили название вторичных мессенджеров. Ими могут быть различные ионы, циклические нуклеотиды, продукты деградации липидов и целый ряд других химических соединений биогенного происхождения.

Использование эукариотами системы вторичных мессенджеров переводит их на новый уровень интеграции всех метаболических и катаболических процессов, что необходимо для существования многоклеточных организмов. В частности, вторичные мессенджеры позволяют

многократно усиливать первичный регуляторный сигнал от внеклеточных регуляторных молекул, которые благодаря этому осуществляют свое действие, находясь в небольших концентрациях во внеклеточном пространстве. Кроме того, многие группы клеток и тканей приобретают способность к однотипной и одновременной реакции на первичный регуляторный сигнал, например на действие гормона какого-либо органа эндокринной системы. Это обеспечивает возможность быстрой адаптации многоклеточного организма к изменяющимся условиям внутренней и окружающей среды.

Трансмембранный перенос первичных сигналов. Для того чтобы первичный регуляторный сигнал достиг ядра и оказал свое воздействие на экспрессию генов-мишеней, он должен пройти через двухслойную мембрану именно тех клеток, которым он предназначен. Как правило, это достигается благодаря наличию на поверхности клеток рецепторов белковой природы, специфически выбирающих из окружающей среды сигналы, распознать которые они в состоянии (рис. 1.22). В простейшем случае, когда в качестве низкомолекулярных регуляторов выступают гидрофобные химические соединения, растворимые в липидах мембран (например стероидные гормоны), для их переноса не используются рецепторы, и они проникают в клетку путем радиальной диффузии. Внутри клеток такие соединения специфически взаимодействуют с белковыми рецепторами, а образующийся комплекс переносится в ядро, где оказывает свое регуляторное воздействие на транскрипцию соответствующих генов (см. рис. 1.22,а).

В отличие от этого рецепторы мембран, ориентированные во внеклеточное пространство, обладают способностью осуществлять транспорт лиганда-регулятора внутрь клеток посредством эндоцитоза (поглощения путем втягивания мембраны) комплекса лиганд-рецептор в составе мембранных везикул. Такой механизм используется, в частности, для переноса внутрь клеток молекул холестерина, ассоциированных с рецепторами липопротеинов низкой плотности (см. рис. 1.22,б). Другой тип рецепторов, ориентированных на внеклеточные лиганды, – это трансмембранные молекулы или группа молекул. Взаимодействие с лигандом внешней части таких молекул сопровождается индукцией ферментативной активности, ассоциированной с внутриклеточной частью того же самого полипептида (см. рис. 1.22,в). Примерами подобных рецепторов, обладающих активностью тирозиновых протеинкиназ, являются

рецепторы инсулина, эпидермального фактора роста или фактора роста тромбоцитов. В синапсах нейронов и местах контакта нейромышечных тканей лиганды-нейромедиаторы (например ацетилхолин или γ -аминомасляная кислота) взаимодействуют с трансмембранными ионными каналами (см. рис. 1.22,а). В ответ на это происходит открытие ионных каналов, сопровождаемое перемещением ионов через мембрану и быстрым изменением трансмембранного электрического потенциала. Другие трансмембранные рецепторы осуществляют связь белков внеклеточного матрикса с микрофиламентами цитоскелета клеток и регуляцию формы клеток, зависящую от внеклеточного матрикса, их подвижности и роста (см. рис. 1.22,б). Наконец, большая группа внеклеточных сигналов распознается рецепторами, ассоциированными на внутренней поверхности мембраны с GTP-связывающими белками, которые, в свою очередь, в ответ на первичный сигнал начинают синтез вторичных мессенджеров, регулирующих активность внутриклеточных белков (см. рис. 1.22,в). Классификация по структурному признаку рецепторов, осуществляющих перенос сигнала в клетки через мембраны, приведена в табл. 1.12.

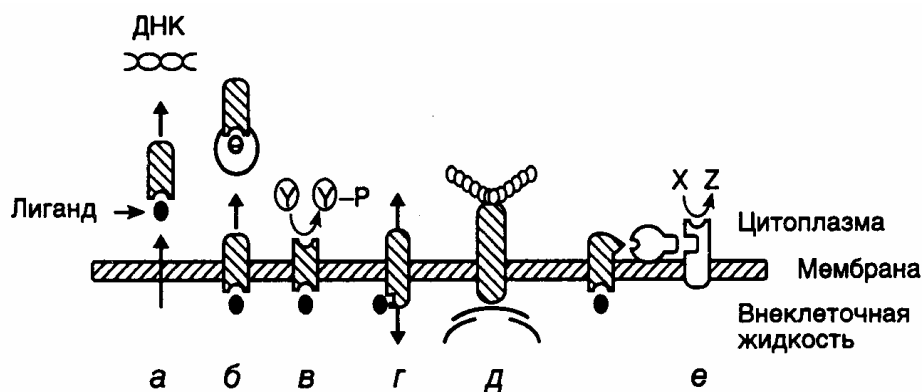


Рис. 1.22. Способы передачи внеклеточных регуляторных сигналов через мембраны эукариотических клеток (а–е)

Y и Y-P – нефосфорилированные и фосфорилированные остатки Тир в белках соответственно. Показано также превращение предшественника X во вторичный мессенджер Z

Все рецепторы, участвующие в трансмембранной передаче сигнала, подразделяют на три класса. При этом, как правило, учитывается сходство или различие вторичных структур субъединиц, а не особенности их аминокислотных

последовательностей. Рецепторы 1-го класса образуют олигомерные структуры вокруг пор в мембранах. Перенос сигнала в этом случае происходит в результате открытия или (в одном случае) закрытия ионных каналов. Основная часть рецепторов 2-го класса погружена в мембраны, и каждая из субъединиц содержит последовательности, распознаваемые G-белками (см. ниже). Для всех субъединиц этого класса характерно наличие трансмембранной (ТМ) последовательности, которая 7 раз пересекает мембрану. Субъединицы рецепторов 3-го класса минимально погружены в мембраны, что обеспечивает подвижность рецепторов и возможность их интернализации (перехода в цитоплазму клеток в составе мембранной везикулы). Большая часть полипептидных цепей этих субъединиц экспонирована наружу клеток.

Вторичные мессенджеры. Гипотеза о том, что действие гормонов на метаболизм клеток и экспрессию генов опосредуется внутриклеточными вторичными мессенджерами, впервые появилась после открытия в конце 1950-х годов Е. Сазерлендом циклического аденозин-3',5'-монофосфата (сАМР). К настоящему времени список вторичных мессенджеров расширился и включает циклический гуанозин-3',5'-монофосфат, фосфоинозитиды, ионы Ca^{2+} и H^+ , метаболиты ретиноевой и арахидоновой кислот, закись азота (NO), а также некоторые другие химические соединения биогенного происхождения. Подробное рассмотрение особенностей механизма действия каждого из них выходит за рамки данной монографии, хотя все они могут оказывать специфическое влияние на экспрессию генов.

Таблица I.12

Рецепторы мембран, осуществляющие трансмембранный перенос сигнала

Класс рецепторов	Четвертичная структура	Система переноса сигнала	Лиганд
1. Олигомеры, окружающие каналы: а) активируемые снаружи; б) активируемые изнутри	Гетеромеры / гомомеры	Ионные каналы, регулируемые медиаторами	а) γ -аминомасляная кислота, Gly, ацетилхолин и т.п. б) cGMP, cAMP, ATP, ионы Ca^{2+} и т.п.
2. Полипептиды с семью гидрофобными доменами. Суперсемейства: I. Основное суперсемейство II. Рецепторы секретина, VIP, паратиреоидного гормона и кальцитонина III. Рецепторы глутамата	Мономеры / гомодимеры посттрансляционно образующиеся гетеродимеры	/ Через G-белки: /, а) вместе с диффундируемым переносчиком; б) непосредственно действуют на каналы в) после расщепления пептидным гормоном, действующим как сайт-специфическая протеиназа с образованием самоактивирующегося рецептора	а) низкомолекулярные медиаторы (кроме Gly): нейропептиды, одоранты, цитокины (IL-8), липиды и подобные агонисты (PAF, эйкозаноиды) б) атриальные мускариновые, нейроновые, α_1 -адренергические лиганды в) тромбин

Таблица I.12 (окончание)

Класс рецепторов	Четвертичная структура	Система переноса сигнала	Лиганд
3. Полипептиды с одним гидрофобным доменом: I. Содержащие одну трансмембранную ТМ-последовательность II. Не содержащие ТМ-последовательность, но обладающие якорной последовательностью для закрепления в мембране	Мономеры / гомодимеры /, гетеродимеры, образующиеся посттрансляционно / нативные гетеродимеры / гетеротримеры	/ С помощью лигандсвязывающей субъединицы, являющейся: а) тирозиновой киназой, стимулируемой лигандами б) гуанилатциклазой, стимулируемой лигандами в) с неизвестной ферментативной активностью	Полипептиды: а) митогенные факторы роста, инсулин б) натрийуретические пептиды в) нейротрофины, гормон роста, пролактин и цитокины

Как было упомянуто выше, внеклеточные сигналы, воспринимаемые рецепторами на поверхности клеток, запускают цепь внутриклеточных биохимических реакций, опосредуемых вторичными мессенджерами, в которые вовлекаются десятки и даже сотни внутриклеточных белков. Для организации адекватного координированного ответа на конкретный внеклеточный сигнал эукариотическая клетка использует две основные стратегии. В соответствии с одной из них происходит изменение активности предсуществующих белков (ферментов, белков цитоскелета, ионных каналов и т.п.) как следствие аллостерических воздействий или в результате ковалентных модификаций (фосфорилирование протеинкиназами или дефосфорилирование). Индуцированные таким образом новые активности белков, в свою очередь, вызывают ответ клетки, основанный на второй стратегии – изменении уровней экспрессии конкретных генов. В результате реализации второй стратегии в клетках меняются число молекул конкретных белков и их качественный состав.

Циклический AMP в роли вторичного мессенджера. В ряде хорошо изученных случаев внеклеточные лиганды после взаимодействия с рецепторами индуцируют образование вторичных мессенджеров через участие GTP-связывающих и GTP-гидролизующих гетеродимерных белков, названных G-белками. Во всех этих системах имеет место последовательность реакций, отображенная на рис. 1.23,а. Внеклеточный лиганд специфически распознается трансмембранным рецептором, который, в свою очередь, активирует соответствующий G-белок, локализованный на цитоплазматической поверхности мембраны. Активированный G-белок изменяет активность эффектора (обычно фермента или белка ионного канала, в рассматриваемом случае – аденилатциклазы), который повышает внутриклеточную концентрацию вторичного мессенджера (в нашем примере – cAMP). Каждый вид рецептора взаимодействует только с определенным представителем семейства G-белков, а каждый G-белок – со специфическим классом эффекторных молекул. Таким образом, в одном конкретном случае гормон или нейромедиатор, реагируя со своим рецептором, вызывает активацию G_s -белка, стимулирующего аденилатциклазу. Этот фермент-эффектор превращает внутриклеточный АТФ в cAMP – классический вторичный мессенджер. Внутриклеточный уровень cAMP может специфически понижаться под действием фосфодиэстеразы, которая превращает cAMP в 5'-AMP.

cAMP активирует множество cAMP-зависимых протеинкиназ, каждая из которых фосфорилирует определенные белки-субстраты. В большинстве клеток животных присутствуют, по крайней мере, две хорошо

охарактеризованные сАМР-зависимые протеинкиназы, фосфорилирующие белки-мишени по остаткам Ser и Thr (серин/треониновые А-киназы). Обе А-киназы представляют собой тетрамеры, состоящие из регуляторного (R) и каталитического (C) димеров полипептидных цепей. R-Димер является мишенью для сАМР, с которым он взаимодействует. Это сопровождается диссоциацией комплекса и освобождением C-цепей, обладающих протеинкиназной активностью. Образующиеся полипептиды, свободно диффундируя в цитоплазме, попадают в ядро, где могут фосфорилировать подходящие белки-мишени, в том числе, факторы транскрипции, что сопровождается их активацией и индукцией транскрипции соответствующих генов. Внутрядерными мишенями киназы А являются, в частности, факторы транскрипции CREB, CREM τ , AP2, SRF, Sp1, участвующие в контроле большого числа клеточных функций, включая пролиферацию и дифференцировку клеток, метаболизм гликогена, регуляцию ионных каналов и т.д.

Каким же образом такой простой механизм обеспечивает избирательность регуляторных воздействий сАМР? Специфичность обеспечивается наличием в клетках определенных типов только им присущих тканеспецифических белков, являющихся субстратами для А-киназ. Например, клетки печени обогащены фосфорилазой-киназой и гликогенсинтазой, активность которых регулируется избирательным фосфорилированием их по сАМР-зависимому механизму, что сопровождается накоплением или освобождением углеводов в гепатоцитах. Адипоциты обогащены липазой, фосфорилирование которой по тому же механизму приводит к освобождению этими клетками свободных жирных кислот. Точно также в клетках других типов, запрограммированных на определенные тканеспецифические функции, содержатся специфические наборы ферментов, активность которых регулируется через их сАМР-зависимое фосфорилирование.

При понижении концентрации гормонов во внеклеточной среде и уменьшении уровня гормонального воздействия на рецепторы внутриклеточное содержание сАМР быстро уменьшается, так как фосфодиэстераза сразу же превращает сАМР в 5'-АМР. Одновременно происходит дефосфорилирование белков-мишеней А-киназ под действием фосфатаз. Активность некоторых фосфатаз также регулируется по сАМР-зависимому механизму. Кроме того, большинство клеток синтезирует белок, названный ингибитором протеинкиназы (PKI), который блокирует активность C-субъединиц А-киназы. Это сопровождается инактивацией соответствующих факторов транскрипции и подавлением экспрессии регулируемых ими генов.

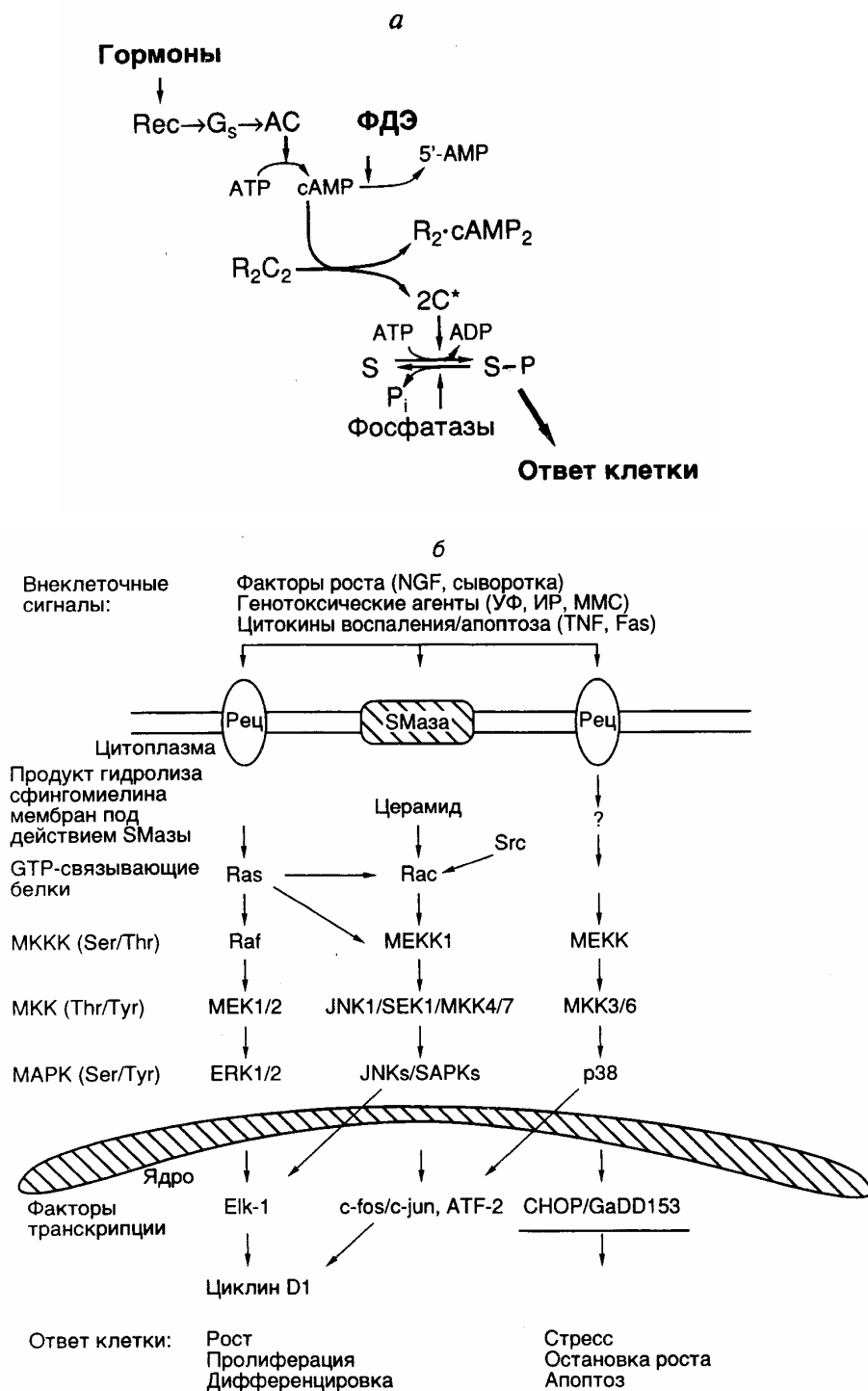


Рис. 1.23. Механизмы передачи сигнала с участием сАМР в качестве вторичного мессенджера (а) и протеинкиназ, активируемых митогенами (МАРК) (б), а также регуляции клеточного цикла (в)
 а: Рес – рецепторы, G_s – G-белок, АС – аденилатциклаза, ФДЭ – фосфодиэстераза, R и C – соответственно регуляторная и каталитическая субъединицы протеинкиназы, S и S-P – белок-субстрат протеинкиназы и его фосфорилированная форма соответственно 2C* – освобожденный

димер каталитической субъединицы А-киназы, P_i – неорганический ортофосфат

б: УФ – ультрафиолетовый свет, ИР – ионизирующая радиация, MMS – метилметансульфонат, SМаза – сфингомиелиназа, MAPKK – киназы, фосфорилирующие MAPK, MAPKKK – киназы, фосфорилирующие MAPKK

в: Образование специфических комплексов циклин–CDK обеспечивает прохождение клетки через соответствующие фазы клеточного цикла.

Отмечены места действия белков-ингибиторов клеточного цикла

Передача сигнала с участием протеинкиназ, активируемых митогенами (МАРК). Протеинкиназы, активируемые митогенами (МАРК – mitogen activated protein kinases), играют исключительно важную роль в регуляции экспрессии генов при всех основных проявлениях жизнедеятельности клеток: их пролиферации и дифференцировке, а также задержке роста и апоптозе в ответ на стрессовые воздействия окружающей среды. После получения внеклеточных сигналов в виде митогенного или генотоксического (мутагенного) воздействия, а также в ответ на действие цитокинов, вызывающих реакции воспаления или апоптоз, в клетках начинают развиваться каскады реакций фосфорилирования, завершающиеся специфической активацией или подавлением активности факторов транскрипции или других регуляторных белков, что сопровождается изменением уровней экспрессии соответствующих генов (см. рис. 1.23,б). МАРК-каскады реакций фосфорилирования протеинкиназ и других регуляторных белков обеспечивают пошаговое декодирование первичных эффекторных сигналов путем их передачи от поверхности клеток к ядру или другим внутриклеточным компонентам, завершающееся кооперативными ответами клеток организма.

По крайней мере, 11 известных МАРК животных осуществляют регуляторное фосфорилирование ядерных факторов транскрипции, белков цитоскелета клетки и белков-участников передачи сигнала на последних этапах этого процесса. К членам семейства МАРК относятся: 1) киназы, регулируемые внеклеточными сигналами, ERK1 и 2 (extracellular signal-regulated kinases); 2) киназы N-концевой части фактора транскрипции Jun и протеинкиназы, активируемые стрессом JNK/SAPK α , β и γ (NH_2 -terminal Jun kinase/stress-activated protein kinases); а также 3) группа МАРК p38, состоящая из четырех белков α , β , γ и δ (см. рис. 1.23,б). МАРК этих групп специфически распознаются и фосфорилируются протеинкиназами 1) MEK1 и 2, известными также под аббревиатурой MKK1 и 2; 2) JNKK1, SEK1, а также MKK4 и 7; 3) MKK3 и 6. Полипептидные цепи МАРК и их киназ МКК обладают высокой гомологией, что

указывает на возможное происхождение генов всего каскада через дубликацию генов модуля MAPK.

Активация MAPK своими МКК происходит по общему механизму через фосфорилирование аминокислотных остатков, находящихся в одинаковом контексте. При этом МКК являются представителями редкого класса протеинкиназ с двойной специфичностью: они могут фосфорилировать как остатки Ser/Thr, так и остатки Tyr.

Сами киназы MAPK (МКК) также активируются через фосфорилирование остатков Ser/Thr киназами киназ MAP-киназ (МККК, или в другом обозначении MAPККК). В отличие от MAPK, каждая из которых распознается и фосфорилируется специфической протеинкиназой (МКК), любая МКК может быть фосфорилирована и активирована несколькими различными МККК, включая белки семейства Raf, MEK-киназы (MEKK), c-Mos и MLK (multilineage protein kinase). Такая неразборчивость МКК в отношении своих активирующих партнеров обеспечивает большое разнообразие путей активации MAPK, начиная с определенных стадий каскада реакций фосфорилирования.

Одни из непосредственных мишеней воздействия сигнала, передаваемого с участием MAPK, протоонкогены *fos* и *jun* кодируют белки, которые являются основными компонентами многосубъединичного фактора транскрипции AP-1 (подробнее о факторах транскрипции см. ниже). В состав фактора входят гомодимеры или гетеродимеры белков семейства Fos (FosB, Fra-1 и Fra-2) и семейства Jun (c-Jun, Jun-B и Jun-D). Фосфорилирование компонентов AP-1 модулирует (увеличивает или уменьшает) активность фактора. Так, фосфорилирование остатков Ser-63 и Ser-73 в полипептидной цепи c-Jun под действием киназы JNK активирует транскрипцию собственного гена после образования гомодимера c-Jun/c-Jun или гетеродимера c-Jun/ATF2. С другой стороны, индукция гена *c-fos* под влиянием митогенов или стресса (например УФ-облучения) опосредована фосфорилированием белка ELK-1, входящего в состав фактора транскрипции TCF (ternary complex factor), который взаимодействует с регуляторной последовательностью SRE промотора этого гена.

Гены, кодирующие белки Fos и Jun, принадлежат к семейству непосредственно ранних генов, индукция которых не требует синтеза белка *de novo* и происходит чрезвычайно быстро в клетках многих типов в ответ на вышеупомянутые внеклеточные и внутриклеточные стимулы. Имеющиеся данные указывают на то, что многокомпонентные факторы транскрипции AP-1, которые представляют собой гомо- и гетеродимеры белков Fos и Jun, играют

ключевую роль в регуляции пролиферации, терминальной дифференцировки и программируемой гибели клеток. Например, гены *fos/jun* индуцируются временно в покоящихся фибробластах в ответ на действие сыворотки. Однако во время дифференцировки миелоидных клеток происходит их стабильная индукция, и уровень транскрипции генов становится максимальным в зрелых клетках, претерпевших терминальную дифференцировку. Все это указывает на возможность участия белков Fos/Jun в инициации и развитии программы терминальной дифференцировки гематопозитических клеток, а также поддержании их дифференцированного состояния. Передача сигнала с участием MAP-киназ играет не менее важную роль и в регуляции клеточного цикла.

Клеточный цикл и его регуляция. Рост и деление клеток являются одними из тех фундаментальных процессов, которые лежат в основе жизни любого организма. Прежде чем совершить деление, клетка должна с высокой точностью копировать свой геном (клеточную ДНК) и подготовить его передачу в дочернюю клетку, а также синтезировать многочисленные высоко- и низкомолекулярные соединения. Повторяющаяся совокупность событий, обеспечивающих деление эукариотических клеток, получила название клеточного цикла. Продолжительность клеточного цикла зависит от типа делящихся клеток. Некоторые клетки, например нейроны человека, после достижения стадии терминальной дифференцировки прекращают свое деление вообще. Клетки легких, почек или печени во взрослом организме начинают делиться лишь в ответ на повреждение соответствующих органов. Клетки некоторых типов, например клетки эпителия кишечника, делятся на протяжении всей жизни человека. Но даже у этих быстро пролиферирующих клеток подготовка к делению занимает ~24 ч.

Фазы клеточного цикла. Активный клеточный цикл эукариотических клеток разделяют на четыре фазы. Наиболее легко обнаруживаемой является стадия непосредственного деления клеток – митоза, при котором конденсированные метафазные хромосомы поровну распределяются между дочерними клетками (М-фаза клеточного цикла – mitosis). Митоз был первой идентифицированной фазой клеточного цикла, а все остальные события, происходящие в клетке между двумя митозами, были названы интерфазой. Развитие исследований на молекулярном уровне позволило выделить в интерфазе стадию синтеза ДНК, получившую название S-фазы (synthesis). Эти две ключевые стадии клеточного цикла не переходят непосредственно одна в другую. После окончания митоза до начала синтеза ДНК имеет место

кажущаяся пауза (gap) в активности клетки – G_1 -фаза клеточного цикла, в которой внутриклеточные синтетические процессы подготавливают репликацию генетического материала. Второй перерыв в видимой активности (фаза G_2) наблюдается после окончания синтеза ДНК перед началом митоза. В фазе G_2 клетка осуществляет контроль за точностью произошедшей редупликации ДНК и исправляет обнаруженные сбои. В ряде случаев выделяют пятую фазу клеточного цикла (G_0), когда после завершения деления клетка не вступает в следующий клеточный цикл и длительное время остается в состоянии покоя. Из этого состояния она может быть выведена внешними стимулирующими (митогенными) воздействиями. Все перечисленные фазы клеточного цикла не имеют четких временных и функциональных границ, отделяющих их друг от друга, однако при переходе от одной фазы к другой происходит упорядоченное переключение синтетических процессов, позволяющее на молекулярном уровне дифференцировать эти внутриклеточные события.

Циклины и циклин-зависимые киназы. Клетки вступают в клеточный цикл и осуществляют синтез ДНК в ответ на внешние митогенные стимулы. Лимфокины (например интерлейкины), цитокины (в частности интерфероны) и полипептидные факторы роста, взаимодействуя со своими рецепторами на поверхности клеток, индуцируют каскад реакций фосфорилирования внутриклеточных белков, сопровождающихся передачей сигнала от поверхности клеток к ядру и индукцией транскрипции соответствующих генов. Одними из первых активируются гены, кодирующие белки циклины, получившие свое название от того, что их внутриклеточная концентрация периодически изменяется по мере прохождения клеток через клеточный цикл, достигая максимума на его определенных стадиях. Циклины являются специфическими активаторами семейства циклин-зависимых протеинкиназ (CDK – cyclin-dependent kinases) – ключевых участников индукции транскрипции генов, контролирующих клеточный цикл. Активация индивидуальной CDK происходит после ее взаимодействия со специфическим циклином, и образование этого комплекса становится возможным после достижения циклином критической концентрации. В ответ на уменьшение внутриклеточной концентрации конкретного циклина происходит обратимая инактивация соответствующей CDK. Некоторые CDK активируются более чем одним циклином. В этом случае группа циклинов, как бы передавая протеинкиназы друг другу, поддерживает их в активированном состоянии длительное время. Такие волны активации CDK возникают на протяжении G_1 - и S-фаз клеточного цикла.

В настоящее время идентифицировано восемь индивидуальных CDK (CDK1–CDK8), часть которых не участвует непосредственно в регуляции клеточного цикла. Для полипептидных цепей всех CDK характерна высокая (до 75%) структурная гомология. Специфичность же их функционирования обеспечивают уникальные сайты связывания соответствующих активирующих циклинов.

В семействе циклинов (циклин А – циклин J) известны, по крайней мере, 14 индивидуальных белков. Некоторые члены семейства составляют подсемейства. Например, подсемейство циклинов D-типа состоит из трех членов: D1, D2 и D3. Общей структурной особенностью всех циклинов является наличие в их полипептидной цепи последовательности из ~100 аминокислотных остатков, получившей название *циклинового бокса*. Циклины относятся к быстро обменивающимся белкам с коротким временем полужизни, которое составляет у циклинов D-типа 15–20 мин. Это обеспечивает динамизм их комплексов с циклинзависимыми киназами. За внутриклеточную деградацию циклинов отвечает N-концевая последовательность аминокислотных остатков, названная *боксом деструкции* (destruction box). При прохождении клеток через клеточный цикл вслед за активацией отдельных CDK по мере необходимости происходит их инактивация. В последнем случае имеет место протеолитическая деградация циклина, находящегося в комплексе с CDK, которая начинается с бокса деструкции.

Сами по себе циклины не могут полностью активировать соответствующие CDK. Для завершения процесса активации должно произойти специфическое фосфорилирование и дефосфорилирование определенных остатков аминокислот в полипептидных цепях этих протеинкиназ. Большую часть таких реакций осуществляет киназа, активирующая CDK (CAK – CDK activating kinase), которая представляет собой комплекс CDK7 с циклином H. Таким образом, CDK становятся способными выполнять свои функции в клеточном цикле лишь после их взаимодействия с соответствующими циклинами и осуществления посттрансляционных модификаций под действием CAK и других аналогичных белков-регуляторов клеточного цикла.

Начало деления эукариотической клетки. В ответ на митогенный стимул клетка, находящаяся в фазе G_0 или ранней G_1 , начинает свое прохождение через клеточный цикл. В результате индукции экспрессии генов циклинов D и E, которые обычно объединяют в группу циклинов G_1 , происходит увеличение их внутриклеточной концентрации. Циклины D1, D2 и D3 образуют комплекс с киназами CDK4 и CDK6. В отличие от циклина D1 два последних

циклина, кроме того, объединяются с CDK2. Функциональные различия между этими тремя циклинами в настоящее время неизвестны, однако имеющиеся данные указывают на достижение ими критических концентраций при разных стадиях развития фазы G_1 . Эти различия специфичны в отношении типа пролиферирующих клеток.

Активация CDK2/4/6 приводит к фосфорилированию белкового продукта гена ретинобластомы pRb и ассоциированных с ним белков p107 и p130. В начале фазы G_1 белок pRb фосфорилирован слабо, что позволяет ему находиться в комплексе с фактором транскрипции E2F, играющим ключевую роль в индукции синтеза ДНК, и блокировать его активность. Полностью фосфорилированная форма pRb освобождает E2F из комплекса, что приводит к активации транскрипции генов, контролирующей репликацию ДНК. (Подробнее о взаимодействии E2F с pRb см. в разделе 3.2.3.) Концентрация D-циклинов возрастает на протяжении фазы G_1 клеточного цикла и достигает максимума значений непосредственно перед началом S-фазы, после чего начинает уменьшаться. Однако в это время pRb еще фосфорилирован не полностью, и фактор E2F остается в комплексе в неактивном состоянии. Фосфорилирование pRb завершается под действием CDK2, активированной циклином E. Внутриклеточная концентрация последнего становится максимальной в момент перехода клеточного цикла от фазы G_1 к S-фазе. Таким образом, комплекс циклин E-CDK2 как бы принимает эстафету от комплексов циклина D с CDK4 и CDK6 и завершает фосфорилирование pRb, сопровождающееся освобождением активного фактора транскрипции E2F. В результате начинается синтез ДНК, то есть клетка вступает в S-фазу клеточного цикла.

Синтез ДНК в S-фазе клеточного цикла. После вступления клетки в S-фазу происходит быстрая деградация циклина E и активация CDK2 циклином A. Циклин E начинает синтезироваться в конце фазы G_1 и его взаимодействие с CDK2 является необходимым условием для вступления клетки в S-фазу и продолжения синтеза ДНК. Этот комплекс активирует синтез ДНК через фосфорилирование белков в областях начала репликации. Сигналом к завершению S-фазы и переходу клетки к фазе G_2 является активация циклином A другой киназы CDK1 с одновременным прекращением активации CDK2. Задержка между окончанием синтеза ДНК и началом митоза (фаза G_2) используется клеткой для контроля полноты и точности произошедшей репликации хромосом. Последовательность событий в этот период точно не известна.

Митоз. Сигнал к началу деления клетки (митоза) исходит от фактора MPF (M phase promoting factor), стимулирующего М-фазу клеточного цикла. MPF представляет собой комплекс киназы CDK1 с активирующими ее циклинами А или В. Складывается впечатление, что комплекс CDK1–циклин А играет более важную роль в завершении S-фазы и подготовке клетки к делению, тогда как комплекс CDK1–циклин В преимущественно осуществляет контроль последовательности событий, связанных с митозом. В настоящее время идентифицировано два циклина В-типа: В1 и В2. Хотя оба циклина, по-видимому, выполняют одинаковые функции, они действуют в разных частях клетки. Так, циклин В1 ассоциирован преимущественно с микротрубочками, тогда как циклин В2 обнаруживают в районе аппарата Гольджи.

Циклины В1 и В2 присутствуют в очень малых концентрациях в фазе G_1 . Их концентрация начинает увеличиваться в конце S- и на протяжении G_2 -фаз, достигая своего максимума во время митоза, что приводит к замещению ими циклина А в комплексе с CDK1. Однако этого оказывается недостаточным для полной активации протеинкиназы. Функциональная компетентность CDK1 достигается после серии ее фосфорилирований и дефосфорилирований по специфическим остаткам аминокислот. Такой тонкий контроль необходим для предотвращения вступления клеток в митоз до полного завершения синтеза ДНК.

Деление клетки начинается только после того, как CDK1, находящаяся в комплексе с циклином В, фосфорилируется по остаткам Thr-14 и Tyr-16 протеинкиназой WEE1, а также по остатку Thr-161 протеинкиназой САК и затем дефосфорилируется по остаткам Thr-14 и Tyr-15 фосфатазой CDC25. Активированная таким образом CDK1 фосфорилирует в ядре структурные белки, в том числе нуклеолин, ядерные ламины и виментин. После этого ядро начинает проходить через цитологически хорошо различимые, но пока недостаточно изученные на молекулярном уровне стадии митоза. Первая стадия митоза – профаза – начинается после того, как CDK1 полностью фосфорилируется, за ней следуют метафаза, анафаза и телофаза, завершающиеся делением клетки – цитокинезом. Следствием этих процессов является правильное распределение реплицированных хромосом, ядерных и цитоплазматических белков, а также других высокомолекулярных и низкомолекулярных соединений в дочерние клетки. После завершения цитокинеза происходит разрушение циклина В, сопровождаемое инактивацией CDK1, что приводит к вступлению клетки в фазу G_1 или G_0 клеточного цикла.

Фаза G₀ клеточного цикла. Клетки некоторых типов на определенных стадиях дифференцировки могут прекращать свое деление, полностью сохраняя свою жизнеспособность. Такое состояние клеток получило название фазы G₀. Клетки, достигшие состояния терминальной дифференцировки, уже не могут выйти из этой фазы. В то же время клетки, для которых характерна чрезвычайно низкая способность к делению, например гепатоциты, могут снова вступать в клеточный цикл после удаления части печени.

Переход клеток в состояние покоя становится возможным благодаря функционированию высокоспецифических ингибиторов клеточного цикла. При участии этих белков клетки могут прекращать пролиферацию в неблагоприятных условиях окружающей среды, при повреждении ДНК или появлении грубых ошибок ее репликации. Такие паузы используются клетками для репарации возникших повреждений.

Ингибиторы клеточного цикла. В клеточном цикле имеются две основные стадии (точки перехода, контрольные точки R – restriction points), на которых могут быть реализованы негативные регуляторные воздействия, останавливающие продвижение клеток через клеточный цикл. Одна из этих стадий контролирует переход клетки к синтезу ДНК, а другая – начало митоза. Имеются и другие регулируемые этапы клеточного цикла.

Переход клеток от одной фазы клеточного цикла к другой контролируется на уровне активации CDK их циклинами с участием ингибиторов циклинзависимых киназ CKI. По мере необходимости эти ингибиторы могут активироваться и блокировать взаимодействие CDK со своими циклинами, а следовательно, и клеточный цикл как таковой. После изменения внешних или внутренних условий клетка может продолжить пролиферацию или вступить на путь апоптоза.

Имеется две группы CKI: белки семейств p21 и INK4 (inhibitor of CDK4), члены которых внутри семейств обладают похожими структурными свойствами. Семейство ингибиторов p21 включает в себя три белка: сам p21, p27 и p57. Поскольку эти белки были описаны независимо несколькими группами, до сих пор используются их альтернативные названия. Так, белок p21 известен также под именами WAF1 (wild-type p53 activated fragment 1), CIP1 (CDK2 interacting protein 1), SDI1 (senescent derived inhibitor 1) и mda-6 (melanoma differentiation associated gene). Синонимами p27 и p57 являются соответственно KIP1 и KIP2 (kinase inhibiting proteins 1 and 2). Все эти белки обладают широкой специфичностью действия и могут ингибировать различные CDK. В отличие от этого группа ингибиторов INK4 более специфична. В нее входят четыре белка:

p15^{INK4B}, p16^{INK4A}, p18^{INK4C} и p19^{INK4D}. До недавнего времени предполагалось, что все ингибиторы семейства INK4 функционируют во время фазы G₁ клеточного цикла, подавляя активность киназы CDK4. Однако обнаруженный недавно второй белковый продукт гена INK4A – p19^{ARF}, взаимодействует с регуляторным фактором MDM2 белка p53 и инактивирует фактор. Это сопровождается увеличением стабильности белка p53 и остановкой клеточного цикла.

Механизмы контроля перехода от G₁- к S-фазе клеточного цикла. До начала активного клеточного цикла белок p27, находясь в высокой концентрации, предотвращает активацию протеинкиназ CDK4 или CDK6 циклинами D1, D2 или D3. В таких условиях клетка остается в фазе G₀ или ранней фазе G₁ до получения митогенного стимула. После адекватной стимуляции происходит уменьшение концентрации ингибитора p27 на фоне возрастания внутриклеточного содержания циклинов D. Это сопровождается активацией CDK и, в конечном счете, фосфорилированием белка pRb, освобождением связанного с ним фактора транскрипции E2F и активацией транскрипции соответствующих генов.

На этих ранних стадиях фазы G₁ клеточного цикла концентрация белка p27 все еще остается довольно высокой. Поэтому после прекращения митогенной стимуляции клеток содержание этого белка быстро восстанавливается до критического уровня и дальнейшее прохождение клеток через клеточный цикл блокируется на соответствующем этапе G₁. Эта обратимость возможна до тех пор, пока фаза G₁ в своем развитии не достигает определенной стадии, называемой точкой перехода, после прохождения которой клетка становится коммитированной к делению, и удаление факторов роста из окружающей среды не сопровождается ингибированием клеточного цикла. Хотя с этого момента клетки становятся независимыми от внешних сигналов к делению, они сохраняют способность к самоконтролю клеточного цикла.

Ингибиторы CDK семейства INK4 (p15, p16, p18 и p19) специфически взаимодействуют с киназами CDK4 и CDK6. Белки p15 и p16 идентифицированы как супрессоры опухолевого роста, и их синтез регулируется белком pRb. Все четыре белка блокируют активацию CDK4 и CDK6, либо ослабляя их взаимодействие с циклинами, либо вытесняя их из комплекса. Хотя оба белка p16 и p27 обладают способностью ингибировать активность CDK4 и CDK6, первый имеет большее сродство к этим протеинкиназам. Считается, что если концентрация p16 повышается до уровня,

при котором он полностью подавляет активность киназ CDK4/6, белок p27 становится основным ингибитором киназы CDK2.

На ранних стадиях клеточного цикла здоровые клетки могут распознавать повреждения ДНК и реагировать на них задержкой прохождения клеточного цикла в фазе G_1 до репарации повреждений. Например, в ответ на повреждения ДНК, вызванные ультрафиолетовым светом или ионизирующей радиацией, белок p53 индуцирует транскрипцию гена белка p21. Повышение его внутриклеточной концентрации блокирует активацию CDK2 циклинами E или A. Это останавливает клетки в поздней фазе G_1 или ранней S-фазе клеточного цикла. В это время клетка сама определяет свою дальнейшую судьбу – если повреждения не могут быть устранены, она вступает в апоптоз, т.е. совершает самоубийство.

Регуляция перехода клеточного цикла от фазы G_2 к фазе М. Ответ клетки на повреждения ДНК может наступить и позднее – перед началом митоза. И в этом случае белок p53 индуцирует синтез ингибитора p21, который предотвращает активацию киназы CDK1 циклином B и задерживает дальнейшее развитие клеточного цикла. Само прохождение клетки через митоз также жестко контролируется – последующие стадии не начинаются без полного завершения предыдущих. Некоторые из этих ингибиторов были идентифицированы у дрожжей, но их гомологи у животных пока остаются неизвестными. Например, недавно описаны два белка дрожжей BUB1 (budding uninhibited by benomyl) и MAD2 (mitotic arrest deficient), которые контролируют присоединение конденсированных хромосом к митотическому веретену в метафазе митоза. До завершения правильной сборки этих комплексов белок MAD2 образует комплекс с протеинкиназой CDC20 и инактивирует ее. CDC20 после активации фосфорилирует белки и в результате блокирует те их функции, которые препятствуют расхождению каждой из двух гомологичных хроматид во время цитокинеза.

3.2.2. Механизмы позитивной регуляции транскрипции

При обсуждении механизмов внутриклеточной передачи сигнала были упомянуты регуляторные белки, взаимодействующие со специфическими последовательностями нуклеотидов генов и получившие название факторов транскрипции. Именно эти белки вносят основной вклад в тонко регулируемую экспрессию генов на уровне транскрипции. Больше тысячи таких белков разных организмов уже идентифицированы, а общее их число, по некоторым оценкам,

составляет не менее нескольких тысяч. В табл. I.13 суммирована информация о 996 известных факторах транскрипции, представленная в базе данных TRANSFAC к концу 1999 г. в соответствии с классификацией, принятой в той же базе данных. Эта классификация основана на гомологии первичных и вторичных структур факторов транскрипции. Каждому фактору соответствует пятизначный цифровой код, первая цифра которого относится к суперклассу, вторая – к классу, третья – к семейству, четвертая – к подсемейству и, наконец, пятая является видовой характеристикой фактора.

Все известные факторы транскрипции на основании гомологии первичных структур их полипептидных цепей разделяют на четыре больших суперкласса: 1) факторы с доменами, обогащенными основными аминокислотами (284 фактора); 2) факторы с ДНК-связывающими доменами, координирующими ионы Zn^{2+} (148 факторов); 3) факторы, содержащие домены типа "спираль–поворот–спираль" (369 факторов); 4) факторы, содержащие домены типа β -scaffold, образующие контакты с малой бороздкой ДНК (156 факторов). Кроме того, 39 факторов, не попадающих по своим структурным особенностям ни в один из вышеперечисленных суперклассов, отнесены к суперклассу 0. Внутри суперклассов факторы разделяются на классы структурно родственных белков, которые, в свою очередь, подразделяются на семейства и подсемейства. Такое разнообразие регуляторных белков у эукариот обусловлено наличием у них большого количества генов и необходимостью тонкой регуляции их экспрессии на протяжении жизненного цикла.

Таблица I.13

Классификация факторов транскрипции

Фактор транскрипции	Характеристика
Суперкласс 1: факторы с основными доменами (basic domains) Класс 1.1: факторы с доменами типа "лейциновая застежка" (leucine zipper) (bZIP)	К ДНК-связывающему участку полипептидной цепи, обогащенному основными аминокислотными остатками, примыкает домен типа "лейциновая застежка", в котором каждым седьмым аминокислотным остатком является Leu. Этот домен обеспечивает димеризацию факторов, необходимую для их взаимодействия с ДНК. Все боковые остатки Leu выступают с одной стороны α-спирали и при воздействии двух полипептидных цепей переплетаются друг с другом, как в застежке-молнии, что обеспечивает специфичность димеризации. Во время взаимодействия димера с ДНК основные участки полипептидных цепей также приобретают α-спиральную конформацию. Предполагают, что два α-спиральных участка А и В, разделенные точкой перегиба, при связывании с ДНК захватывают ее наподобие ножниц. При этом ДНК контактирует с полипептидными цепями своей большой бороздкой на протяжении одного витка спирали. В классе семь семейств: 1) AP-1-подобные факторы (включает семь подсемейств: Jun, Fos, Maf, NF-E2, AP-1-подобные факторы грибов, CRE-BP/ATF и прочие факторы; всего 47 факторов); 2) CREB (CREB, ATF-1, CREM, BBF-2, dCREB2, SCO1, HAC1, Pcr1; всего 31 фактор); 3) C/EBP-подобные факторы (C/EBPα-ε, CHOP-10; всего 13 факторов); 4) bZIP/PAR (DBP, VBP, Hlf, TEF; всего четыре фактора); 5) G-бокс-связывающие факторы растений (пять подсемейств: CPRF-2, EmBP-1, HBP-1a, TGA-1a, TGA-1b; всего 24 фактора); 6) только ZIP (SWI6, STE4; всего пять факторов); 7) другие факторы bZIP (Giant, OPI1). Общее число известных факторов в классе – 126.
Класс 1.2: факторы с доменами типа "спираль–петля–спираль" (helix–loop–helix) (bHLH)	К ДНК-связывающему участку полипептидной цепи, обогащенному основными аминокислотными остатками, примыкают две амфипатические α-спирали (см. ниже), соединенные друг с другом полипептидными петлями разной длины. Мотив "спираль–петля–спираль" (HLH) обеспечивает димеризацию факторов и их взаимодействие с ДНК, а область факторов, обогащенная основными аминокислотами, определяет специфичность ДНК-белкового взаимодействия. Во время димеризации α-спирализация полипептидных цепей факторов может значительно усиливаться, а после связывания димеров с ДНК основные домены также α-спирализуются. В димере все четыре HLH-домена, по-видимому, ориентированы параллельно друг другу. В классе девять семейств: 1) повсеместно распространенные факторы (класс А) (E2A, HEB/SCBP,

Фактор транскрипции	Характеристика
Класс 1.3: факторы с двойными доменами спираль–петля–спираль/лейциновая застёжка (bHLH–ZIP)	Daughterless; всего девять факторов); 2) миогенные факторы транскрипции (MyoD, Myogenin, Myf-5; всего восемь факторов); 3) Achaete–Scute (Scute, Achaete, MASH-1; всего девять факторов); 4) Tal/Twist/Atonal/Hen (подсемейства: лимфоидные факторы, мезодермальные Twist-подобные факторы, HEN, Atonal, панкреатические факторы; всего 27 факторов); 5) Hairy (подсемейства: Hairy, Esp, регуляторы грибов; всего 12 факторов); 6) факторы с PAS-доменом (Ahr, Arnt; всего четыре фактора); 7) INO (INO2, INO4); 8) факторы, содержащие только HLH-домен (Ets, Id1, Olf-1; всего девять факторов); 9) другие bHLH-факторы (Deliah, Lc, CBF1). Общее число известных факторов в классе – 82. ДНК-связывающий участок обогащен остатками основных аминокислот, к которому примыкают две амфипатические α-спирали, соединенные полипептидной петлей, а также домен типа "лейциновая застёжка". HLH-мотив обеспечивает димеризацию факторов, необходимую для их взаимодействия с ДНК. В димеризации также участвуют остатки Leu мотива "застёжка". У факторов, связавшихся с ДНК, значительно усиливается α-спирализация полипептидных цепей. В димерах HLH-домены ориентированы параллельно друг другу. Область, обогащенная основными аминокислотами, обеспечивает специфичность взаимодействия факторов с ДНК. В классе два семейства: 1) повсеместно распространенные факторы bHLH–ZIP (подсемейства: TFE3, USF, SREBP, AP-4; всего 16 факторов); 2) факторы, контролирующие клеточный цикл (Myc, Mad/Max, E2F, DRTF; всего 26 факторов). Общее число известных факторов в классе – 42.
Класс 1.4: NF-1	В классе одно семейство ядерных факторов NF-1 (22 фактора).
Класс 1.5: RF-X	Фактор RF-X1 идентифицирован как трансактиватор энхансера I вируса гепатита В, входит в состав семейства гомодимерных и гетеродимерных факторов транскрипции. В классе одно семейство RF-X, включающее 5 представителей: RF-X1–RF-X5
Класс 1.6: bHSH Суперкласс 2: факторы с ДНК-связывающими доменами, содержащими ион цинка в координационном центре Класс 2.1: факторы типа	В классе одно семейство AP-2 (AP-2α, AP-2β, AP-2γ). Общее число известных факторов в классе – семь. Все факторы содержат домен "цинковые пальцы", характерный для ядерных рецепторов. По два таких

Фактор транскрипции	Характеристика
ядерных рецепторов, содержащие мотив "цинковые пальцы" Cys4	домена различного размера, аминокислотного состава и механизма действия присутствуют в каждой молекуле рецептора. Каждый палец содержит четыре остатка Cys, координирующих ион Zn^{2+}. Половина домена, включающая вторую пару остатков Cys, образует α-спираль. Спираль первого пальца взаимодействует с большой бороздкой ДНК. Последовательность между первыми двумя остатками Cys обеспечивает димеризацию факторов, связавшихся с ДНК. В классе два семейства: 1) рецепторы стероидных гормонов (подсемейства: рецепторы кортикоидов (GR), прогестерона (PR), андрогенов (AR), эстрогенов (ER); всего десять факторов); 2) факторы, подобные рецепторам тиреоидных гормонов (подсемейства: рецепторы ретиноевой кислоты (RAR), ретиноида X (RXR), тиреоидных гормонов (T3R), витамина D (VDR), NGFI-B, FTZ-F1, PPAR, EcR, ROR, TII/COUP, HNF-4, CFI, Knirops; всего 69 факторов).
Класс 2.2: факторы с разными доменами типа "цинковые пальцы" Cys4	Каждая молекула факторов содержит по два домена типа "цинковые пальцы" GATA-типа. Каждый палец включает четыре остатка Cys, по-видимому, координирующих один ион Zn^{2+}, однако хелатные агенты не подавляют взаимодействие факторов с ДНК. В классе три семейства: 1) GATA-факторы (подсемейства: GATA-факторы позвоночных и метаболические регуляторы грибов, всего девять факторов); 2) Trithorax (Ttx, Hrx); 3) прочие факторы (BUF2); всего 12 факторов.
Класс 2.3: факторы с доменами типа "цинковые пальцы" Cys2His2	Все факторы содержат мотив "цинковые пальцы" типа TFIIIA/Krueppel, в котором каждый палец включает по два остатка Cys и His, координирующих ион Zn^{2+}. В некоторых случаях один остаток His может быть заменен на остаток Cys. Ион Zn^{2+} необходим для взаимодействия факторов с ДНК. Первая половина аминокислотной последовательности пальца обычно образует два антипараллельных β-слоя, а вторая половина организована в α-спираль и частично 3–10-спираль. Гидрофобные связи между β-слоями и α-спиралями образуются при участии консервативных остатков Phe и Leu. Факторы взаимодействуют с большой бороздкой ДНК α-спиральными участками полипептидных цепей. Последовательность аминокислот, связывающая пальцы, чаще всего TGEKPI. В классе пять семейств: 1) повсеместно распространенные факторы (TFIIIA, Sp1, YY1; всего пять факторов); 2) регуляторы развития или клеточного цикла (подсемейства: Egr/Krox, Kruppel-подобные факторы, GLI-подобные факторы и прочие факторы; всего 36 факторов); 3) метаболические регуляторы грибов (CE2, CreA; всего шесть факторов); 4) большие факторы с NF-6B-подобными ДНК-связывающими свойствами (HIV-EP1, αA-CRYBP1, AGIE-BP1; всего шесть факторов); 5) регуляторы вирусов (T-Ag). Всего в классе 55 факторов.
Класс 2.4: факторы,	Все факторы содержат кластер из шести остатков Cys, координирующих два иона Zn^{2+}. Это означает,

Фактор транскрипции	Характеристика
содержащие кластер из шести остатков Cys, координирующих два иона Zn^{2+}	что две из шести SH-групп координируют по два иона Zn^{2+} каждая. Часто встречаются у грибов. В классе одно семейство: метаболические регуляторы грибов (ARG RII, GAL4, UGA3; всего 11 факторов).
Класс 2.5: факторы, содержащие домены типа "цинковые пальцы" переменного состава Суперкласс 3: факторы с доменами типа "спираль–поворот–спираль"	В классе два семейства: 1) факторы с доменами типа "цинковые пальцы" Cx7Hx8Cx4 (BAF1); 2) факторы с доменами типа "цинковые пальцы" Cx2Hx4Hx4C (Bir3). Всего в классе два фактора.
Класс 3.1: факторы, содержащие гомеодомены	Все факторы содержат в полипептидных цепях три последовательно расположенных α -спиральных участка. Спираль 3 контактирует преимущественно с большой бороздкой ДНК, однако обнаружены контакты и с малой бороздкой. Пространственное расположение спиралей 2 и 3 напоминает таковое у регуляторов транскрипции прокариот. В классе четыре семейства: 1) факторы, содержащие только гомеодомены (подсемейства: AbdB, Antp, Cad, Cut, Dll, Ems, En, Eve, Prd, HD-ZIP, H2.0, HNF1, Lab, Msh, NK-2, Bcd, XANF, PBS, неопределенное; всего 158 факторов); 2) факторы, содержащие POU-домены (подсемейства: I, II, III, IV, V, VI, другие POU-факторы; всего 58 факторов); 3) факторы, содержащие гомеодомены с LIM-областью (подсемейства: факторы с гомеодоменами, содержащими LIM-область; (ко)факторы с гомеодоменами, содержащими только LIM-область, всего 15 факторов); 4) факторы с гомеодоменами и мотивами типа "цинковые пальцы" (представители: ATBF1, Zfh1; всего четыре фактора). Общее число известных факторов в классе – 235.
Класс 3.2: факторы, содержащие парный бокс	Все факторы содержат ДНК-связывающий домен длиной в ~130 аминокислотных остатков. Как правило, его N-концевая половина обогащена основными аминокислотами, а С-концевая половина сильно поляризована. Домен, по-видимому, содержит три α -спиральных участка. В классе два семейства: 1) факторы, содержащие парный бокс и гомеодомен (Prd, Pax-3; всего семь факторов); 2) факторы, содержащие только парный бокс (Pax-1, Pax-2; всего девять факторов).
Класс 3.3: факторы, содержащие домены типа "Fork head/winged helix"	Сгруппированы на основании гомологии с факторами транскрипции HNF-3 мышей и Fkh дрозофилы. Домен составлен из ~110 аминокислотных остатков, которые компактно организованы в три α -спиральных участка. Третья α -спираль контактирует с большой бороздкой ДНК. Домен образует также

Фактор транскрипции	Характеристика
Класс 3.4: факторы теплового шока	несколько контактов с малой бороздкой. Взаимодействие факторов с ДНК приводит к ее изгибу под углом $\sim 13^\circ$. В классе три семейства: 1) регуляторы развития (представители: Fkh, Slp1, всего 11 факторов); 2) тканеспецифические регуляторы (представители: HNF-3, SGF-1, всего пять факторов); 3) другие регуляторы (представители: ILF, FKHR, HTLF, всего 26 факторов). Взаимодействуют с ДНК, организованной в нуклеосомы, только в присутствии основного фактора транскрипции TBP. В классе одно семейство HSF (девять факторов).
Класс 3.5: факторы, содержащие кластеры Trp	Полипептидные цепи факторов содержат кластеры смежных триптофановых остатков. Кластеры разделены участками цепи длиной 12–21 остаток (в семействе Myb 19–21 остаток). В классе три семейства: 1) Myb (Myb и Myb-подобные факторы соответственно 15 и пять факторов); 2) факторы Ets-типа (c-Ets-1, Erg-1; всего 36 факторов); 3) факторы, регулируемые интерфероном (IRF-1, Pp; всего семь факторов). Всего в классе 63 фактора.
Класс 3.6: факторы, содержащие TEA-домены	TEA-домен полипептидной цепи, содержит три α -спирали. Между спиралями 1 и 2 расположены 16–18 остатков аминокислот, а между спиралями 2 и 3 – три–восемь остатков, с неупорядоченной пространственной структурой. Установлено, что указанный домен фактора TEF-1 взаимодействует с ДНК и, по крайней мере, еще два участка полипептидной цепи создают контакты фактора с матрицей. В классе одно семейство TEA (представители: TEF-1, Sd, всего четыре фактора).
Суперкласс 4: факторы с доменами β -Scaffold, контактирующие с малой бороздкой ДНК	
Класс 4.1: факторы, содержащие RHR-домен (Rel homology region)	ДНК-связывающий домен Rel-типа содержит два субдомена, каждый из которых организован в два β -цилиндра с пятью петлями, участвующими в контактах с большой бороздкой ДНК. Наибольший вклад в специфичность взаимодействия вносит первая петля N-концевого субдомена, которая сильно консервативна. Участие полипептидных петель в контактах белок–ДНК обеспечивает факторам возможность взаимодействия с разными нуклеотидными последовательностями. Дополнительные взаимодействия осуществляются через α -спиральные участки N-концевых частей факторов, которые образуют прочные контакты с малой бороздкой ДНК. Второй C-концевой домен обеспечивает главным образом димеризацию белков.

Фактор транскрипции	Характеристика
Класс 4.2: p53	В классе 3 семейства: 1) Rel/ankyrin-факторы (NF-κB1, RelA; всего 11 факторов); 2) только ankyrin-факторы (IκBα, Bcl-3; всего 8 факторов), 3) NF-AT (6 факторов). Общее число факторов – 25. ДНК-связывающий домен организован в виде β-сэндвича, который является остовом для двух больших петель и мотива типа "петля–слой–спираль". N-Концевые участки трех петель (остатки 112–124) образуют контакты с большой бороздкой ДНК в каноническом сайте связывания, тогда как С-концевая петля (остатки 236–251) связывают малую бороздку в АТ-богатом центре сайта связывания. Остаток Arg-248 С-концевой части ДНК-связывающего домена, который чаще всего изменяется под действием мутаций, образует дополнительные контакты с малой бороздкой. Ион Zn^{2+} связывается остатками Cys-176 и His-179, находящимися в составе мотива "петля–спираль", а также остатками Cys-238 и Cys-242 (в составе С-концевой петли). В классе одно семейство p53 (p53 и p53as).
Класс 4.3: факторы, содержащие MADS-бокс	ДНК-связывающие домены обеспечивают димеризацию факторов. В димере, связанном с ДНК, как это показано для фактора SRF, две амфипатические α-спирали образуют соленоид (coiled coil) и ориентированы вдоль малой бороздки ДНК. Эти спирали формируют контакты с большой и малой бороздками. В месте изгиба ДНК обернута вокруг белка. В центре изгиба малая бороздка сжата, а по бокам большие бороздки растянуты. В классе три семейства: 1) регуляторы дифференцировки (подсемейства: MEF-2A-D, факторы гомеотических генов, регуляторы дрожжей; всего 53 фактора); 2) респондеры внешних сигналов (два представителя: SRF, RLM1); 3) регуляторы метаболизма (один представитель: ARG RI). Общее число факторов – 57.
Класс 4.4: факторы, содержащие β -цилиндр (barrel) и α -спиральные участки	В классе одно семейство E2 (E2 и EBNA-1).
Класс 4.5: ТАТА-связывающие белки	Для полипептидов характерны консервативный С-концевой и дивергировавший N-концевой участки. Последний необходим для взаимодействия с промоторами Pol II и Pol III, содержащими ТАТА-бокс, но не с промоторами Pol I, в которых эта последовательность отсутствует. Может связывать ДНК в виде мономера, сгибая молекулу; входит в состав различных TAF.

Фактор транскрипции	Характеристика
Класс 4.6: HMG-факторы	В классе одно семейство: TBP (с одним представителем TBP). Белки этого класса обладают гомологией с хромосомными белками, принадлежащими к группе негистоновых белков с высокой электрофоретической подвижностью (HMG). Участок гомологии содержит ДНК-связывающий домен, который в одних случаях, как и у HMG1, обеспечивает белкам неспецифическое взаимодействие с ДНК, а в других (например у фактора LEF-1) – сайт-специфическое. Домен обнаруживает типичную конформацию в форме буквы L с тремя α-спиралями. Третья спираль вместе с N-концевой частью спирали 1 далеко выдвинуты и образуют длинную часть L-образного домена, а C-концевая часть спирали 1 и спираль 2 – короткую. Взаимодействие с малой бороздкой ДНК вызывает резкий (более чем на 90°) ее изгиб в сторону от белка. Общая топология комплекса напоминает таковую комплекса TBP–TATA-бокс. В классе шесть семейств: 1) SOX (SRY, Sox; всего 15 факторов); 2) TCF-1 (всего 12 факторов); 3) HMG2-подобные факторы (SSRP1, Ixr1; всего четыре фактора); 4) UBF (три фактора); 5) MATA (один фактор: mat-Mc); 6) прочие факторы, содержащие HMG-бокс (всего три фактора). Общее число факторов – 38.
Класс 4.7: гетеромерные CCAAT-факторы	Включает одно семейство: гетеромерные CCAAT-факторы (CP1A/HAP3, CBF-C/HAP-4). Общее число факторов – шесть.
Класс 4.8: Grainyhead	Включает одно семейство Grainyhead. Всего в классе три фактора: CP2, LBP-1a и Grainyhead
Класс 4.9: факторы с доменом холодового шока	Обладают протамино-подобным доменом: последовательностью аминокислот с большим суммарным положительным зарядом, содержащей неупорядоченные остатки Pro. Включает одно семейство csd с семью факторами (DbpA–B, FRG, YB1–3 и др.).
Класс 4.10: Runt	Белки этого класса обладают гомологией с характерным участком полипептидной цепи белка Runt дрозофилы. Домен Runt является частью ДНК-связывающего домена этих факторов. Он сформирован в основном β-слоями, которые не содержат α-спиральных участков, и напоминает домен типа пальмовой ветви, обнаруженный у ДНК-полимеразы β крыс. В классе одно семейство: Runt. (PEBP2α, Runt; всего 15 факторов).
Суперкласс 0: прочие факторы	

Фактор транскрипции	Характеристика
транскрипции Класс 0.1: медь-связывающие белки	Медь-зависимый активатор гена металлотионеина (CUP1) у дрожжей. Не содержит доменов типа "цинковые пальцы". В классе одно семейство: регуляторы грибов (всего два фактора: ACE1/CUP2 и AMT1).
Класс 0.2: HMGI(Y)	Вспомогательные факторы для других факторов транскрипции, таких как NF-κB или ATF-2. Облегчают их включение в транскрипционный комплекс. В классе одно семейство: HMGI (Y), включающее четыре фактора.
Класс 0.3: STAT	Содержат домен, подобный лейциновой застёжке, а также ДНК-связывающий домен, определяющий специфичность взаимодействия факторов с ДНК. Содержат домены SH2 и SH3, первый из которых необходим для гомо- и гетеродимеризации факторов, фосфорилированных по остатку Тир. Опосредуют перенос сигналов фосфорилирования к генам, индуцируемым интерфероном. В классе одно семейство: STAT, включающее восемь факторов.
Класс 0.4: факторы, содержащие домен Pocket	В классе два семейства: 1) Rb (Rb – продукт антионкогена ретинобластомы и p107); 2) CBP (один представитель).
Класс 0.5: E1A-подобные факторы	В классе одно семейство: E1A (два представителя).
Класс 0.6: AP2/EREBP-подобные факторы	В классе три семейства: AP2 (четыре фактора); EREBP (ABI4, EBP, ERF1, TINI, всего 14 представителей); AP2/B3 (RAV1 и RAV2).

Ниже будут подробнее рассмотрены особенности структуры некоторых наиболее часто встречающихся структурно-функциональных доменов факторов транскрипции и механизмы их регуляторного действия на транскрипцию генов.

Факторы транскрипции в роли позитивных регуляторов (активаторов) синтеза РНК. Изучение механизмов функционирования белковых факторов транскрипции – задача непростая. Прежде всего, это связано с их малой концентрацией в клетках и большими сложностями получения факторов в очищенном состоянии для дальнейших биохимических исследований. Развитие генно-инженерных методов позволило получать факторы транскрипции в неограниченных количествах, что не замедлило принести плоды в виде новой информации (подробнее о методах клонирования генов см. в главе 7).

Для того чтобы фактор транскрипции оказал специфическое действие на синтез РНК определенным геном, он, прежде всего, должен распознать этот ген и связаться с определенной последовательностью ДНК. Затем фактор транскрипции должен взаимодействовать с другими факторами или непосредственно с самой РНК-полимеразой для стимуляции или подавления транскрипции на этом гене. Кроме того, необходимость включения и выключения транскрипции в строго определенных месте и времени (тканеспецифический характер экспрессии генов на разных стадиях онтогенеза организма) предполагает наличие механизмов контроля биосинтеза или активации самих факторов транскрипции для упорядочивания их функционирования. Ниже будут рассмотрены механизмы, обеспечивающие эти три этапа функционирования факторов транскрипции – связывание с ДНК, влияние на процесс транскрипции и регуляция их собственной активности.

Механизмы взаимодействия факторов транскрипции с ДНК. Клонирование генов факторов транскрипции (и их фрагментов), а также выделение соответствующих рекомбинантных белков позволили идентифицировать участки их полипептидных цепей, обеспечивающие специфическое взаимодействие факторов с ДНК. Очищенные фрагменты белков были исследованы на способность взаимодействовать с определенными последовательностями ДНК. Это позволило обнаружить несколько структурных элементов (доменов) полипептидных цепей, общих для факторов транскрипции разных типов, которые легли в основу их современной

классификации.

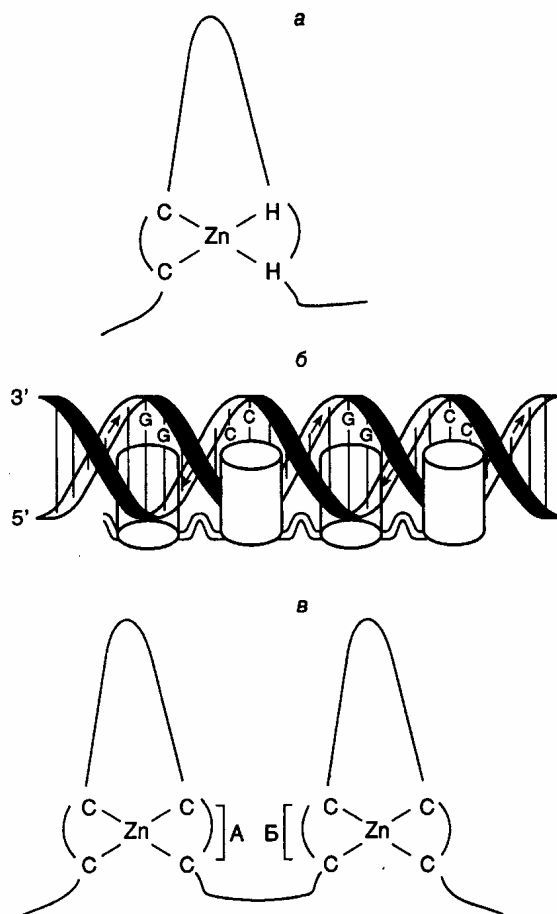


Рис. 1.24. Особенности структуры полипептидных доменов типа "цинковые пальцы" и их взаимодействие с ДНК

а – схема строения домена, содержащего Zn-связывающие остатки Cys (C) и Hys (H); б – предполагаемый механизм взаимодействия доменов типа "цинковые пальцы", изображенных в виде цилиндров, с ДНК. Стрелки указывают полярность цепей ДНК 5'→3'; в – схема строения доменов типа "цинковые пальцы", содержащих только Zn-связывающие остатки Cys. Квадратными скобками отмечены участки полипептидных цепей, участвующие в распознавании палиндромных последовательностей ДНК (А и Б)

Домены типа "цинковые пальцы". Одним из первых факторов транскрипции, полученных в виде очищенного рекомбинантного белка, был фактор TFIIIA, который играет ключевую роль в транскрипции генов 5S рРНК РНК-полимеразой III. ДНК-связывающий участок полипептидной цепи этого фактора содержит девять 30-звенных повторов: Tyr/Phe-Xaa-Cys-Xaa-Cys-

$\text{Xaa}_{2,4}\text{-Cys-Xaa}_3\text{-Phe-Xaa}_5\text{-Leu-Xaa}_2\text{-His-Xaa}_{3,4}\text{-His-Xaa}_5$, где Xaa – остаток любой аминокислоты. Таким образом, каждый из повторов содержит две строго консервативные пары остатков Cys и His, которые взаимодействуют с одним ионом цинка. Это приводит к образованию в свернутой полипептидной цепи пространственной структуры, сформированной консервативными остатками Phe, Leu и несколькими остатками основных аминокислот. Структура выступает над поверхностью белковой глобулы в виде "пальца" (рис. 1.24,а). Вершины пальцев непосредственно контактируют с большой бороздкой ДНК, причем соседние пальцеобразные структуры связываются противоположными сторонами спирали ДНК (см. рис. 1.24,б).

Домены типа "цинковые пальцы" обнаружены у многих факторов транскрипции, обеспечивающих функционирование РНК-полимеразы II, в том числе у фактора Sp1, Kruppel-белка *Drosophila*, белков ADRI и GAL4 дрожжей и белка аденовируса E1A. Интересно, что точковая мутация в гене Kruppel-белка, приводящая к замене лишь одного из остатков Cys на Ser, что, в свою очередь, предотвращает связывание иона Zn^{2+} , фенотипически проявляется как делеция целого гена этого фактора. На данном основании был сделан вывод о том, что способность связывать ионы Zn^{2+} является критической для проявления ДНК-связывающей активности факторов такого типа.

Аналогичные домены обнаружены в полипептидных цепях семейства рецепторов тиреоидных или стероидных гормонов, которые в комплексе с гормонами после переноса в ядра специфически взаимодействуют с определенными последовательностями ДНК, изменяя уровни транскрипции соответствующих генов-мишеней. Однако в этом случае ДНК-связывающие участки состоят из двух пальцев, каждый из которых содержит четыре остатка Cys, взаимодействующие с ионом Zn^{2+} , вместо двух Cys и двух His (см. рис. 1.24,в); у них также отсутствуют консервативные остатки Phe и Leu. Кроме того, элемент из двух цинковых пальцев встречается в полипептидных цепях таких рецепторов только один раз, тогда как в генах, кодирующих цинковые пальцы с Cys-His, подобный элемент может повторяться от 2 до 37 раз.

Исследование мест связывания различных рецепторов стероидных гормонов с ДНК показало, что они взаимодействуют с гомологичными, но не идентичными регуляторными последовательностями нуклеотидов (подчеркнуты):

Глюкокортикоиды/прогестерон
 Эстрогены
 Тиреоидные гормоны/ретиноевая кислота

GGTACANNNTGTTCT
AGGTCANNNTGACCT
TCAGGTCA---TGACCTGA

где N – любые нуклеотиды, --- – отсутствие нуклеотидов.

Методами направленного мутагенеза установлено, что специфичность взаимодействия этих рецепторов со своими последовательностями определяется небольшим числом аминокислотных остатков. Так, замена всего лишь двух аминокислотных остатков в N-концевой части цинкового пальца рецептора глюкокортикоидов на аминокислотные остатки, обнаруживаемые в том же самом участке рецептора эстрогенов, приводит к тому, что мутантный рецептор начинает взаимодействовать с регуляторными последовательностями, узнаваемыми рецепторами эстрогенов, и активировать гены, контролируемые этими последовательностями. Аналогичные замены пяти аминокислотных остатков во втором пальце рецептора эстрогенов также приводят к изменению его специфичности: он приобретает способность взаимодействовать с регуляторной последовательностью тиреоидных гормонов. Как уже упоминалось выше, оба рецептора узнают одни и те же последовательности нуклеотидов, различающиеся лишь расстояниями между ними. При этом первый цинковый палец играет ключевую роль в узнавании последовательности как таковой, а второй определяет оптимальное для взаимодействия расстояние между двумя половинками этой последовательности.

Факторы транскрипции, содержащие мотив "спираль–поворот–спираль". Другой тип пептидных доменов, специфически распознающих регуляторные последовательности на ДНК, характерен для белковых продуктов гомеотических (гомеозисных) генов, впервые обнаруженных у дрозофилы. Гомеотические гены позвоночных и растений играют ключевую роль в их морфогенезе. Полипептидные цепи белков, кодируемых этими генами, содержат высококонсервативную последовательность длиной в 60 аминокислотных остатков, называемую гомеобоксом, или гомеодоменом, которая определяет специфичность взаимодействия белков с регуляторными последовательностями ДНК. Анализ третичной структуры гомеодоменов

показал, что они образуют структуру типа "спираль–поворот–спираль" (helix–turn–helix), в которой за α -спиральным участком следует β -структура с последующим еще одним α -спиральным участком (рис. 1.25,а). Наличие этой пространственной структуры, существование которой было впервые предсказано на основании гомологии с соответствующими аминокислотными последовательностями бактериальных белков-репрессоров, в настоящее время строго доказано методами ЯМР-спектроскопии, в частности для гомеодомена продукта гена *Antennapedia* дрозофилы.

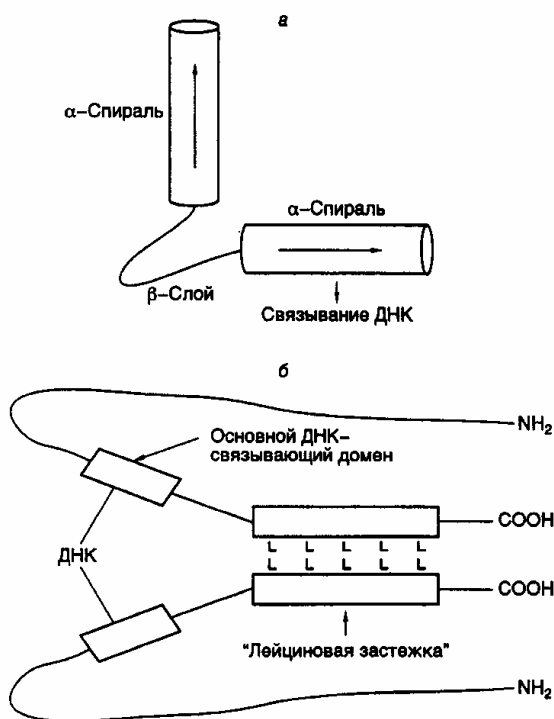


Рис. 1.25. Структуры полипептидных доменов типа "спираль–поворот–спираль" (а) и "лейциновая застежка" (б)
L – остатки Leu

Для бактериальных белков-репрессоров было показано, что при образовании специфических комплексов с ДНК первый α -спиральный участок расположен перпендикулярно большой бороздке ДНК, тогда как второй частично находится в ней и обеспечивает специфические контакты репрессора с операторным участком ДНК. Ключевая роль второго α -спирального участка гомеодомена при изучении специфичности взаимодействия этих белков с ДНК

была определена с помощью мутаций. Так, замена только одного остатка Ser в положении 9 α -спирали белка Prd на Gln, присутствующий в гомеозисном белке Ftz, приводит к связыванию белка Prd с Ftz-сайтами на ДНК.

Позднее была обнаружена большая группа регуляторных белков, в которых гомеобоксы формируют лишь одну часть более протяженного консервативного домена, названного POU-доменом. Эти белки, к которым относятся, в частности, октамерсвязывающие белки (octamer binding proteins) Oct-1 и Oct-2, белок гипофиза Pit-1 и продукт гена unc-86 нематод, используют POU-домены для специфического взаимодействия с ДНК. Относительный вклад аминокислотной последовательности гомеобокса и негомеобоксной POU-последовательности в специфичность связывания белок-ДНК не одинаков для разных белков. Так, в белке Pit-1 специфичность в основном достигается за счет гомеодомена, тогда как в белке Oct-1 основную роль играет другая последовательность POU-домена.

ДНК-связывающий домен типа "лейциновая застежка". Сравнение аминокислотных последовательностей ряда других факторов транскрипции: фактора транскрипции печени C/EBP, дрожжевого фактора GCN4, а также белков-продуктов протоонкогенов Мус, Fos и Jun (протоонкогены – это гены, мутации в которых могут приводить к злокачественному перерождению клеток), позволило выявить еще одну консервативную аминокислотную последовательность, образующую пространственную структуру, названную "лейциновой застежкой" (leucine zipper). В этой структуре остатки Leu, находящиеся в составе α -спирального участка полипептидных цепей факторов, расположены через каждые шесть аминокислотных остатков на равном расстоянии друг от друга и оказываются на одной стороне α -спирали через каждые два витка (см. рис. 1.25,б). Такие домены сами по себе не связываются с ДНК, однако обеспечивают димеризацию содержащих их факторов путем взаимопроникновения лейциновых α -спиралей двух молекул факторов. В результате в димере появляются две правильно расположенные друг относительно друга полипептидные цепи, составленные преимущественно из основных аминокислотных остатков, которые и образуют ДНК-связывающий центр фактора транскрипции.

Димеры белков Fos и Jun связываются с Ap-1-последовательностями

ДНК, что сопровождается активацией соответствующих генов в ответ на воздействие форболовыми эфирами. Однако если белок Jun связывается с Ар-1-последовательностями нуклеотидов в виде гомодимера (белкового комплекса, состоящего из двух идентичных полипептидных цепей), то белок Fos специфически взаимодействует с ДНК только после образования комплекса (гетеродимера) с белком Jun. Эти свойства двух белков являются следствием различий в структуре их лейциновых доменов, которые не позволяют образовываться гомодимеру из полипептидных цепей Fos. Искусственная замена лейцинового домена Fos на соответствующий домен белка Jun обеспечивает условия для получения гомодимера из таких химерных полипептидных цепей. Необходимость образования димера этих двух регуляторных белков перед взаимодействием с ДНК создает дополнительную возможность воздействия на экспрессию соответствующих генов путем контроля над формированием самого белкового комплекса фактора транскрипции.

Основной ДНК-связывающий домен, впервые обнаруженный у белков, содержащих "лейциновую застёжку", позднее был найден и у ряда других белков, регулирующих транскрипцию, в частности у двух белков E12 и E47, которые взаимодействуют с энхансерами иммуноглобулиновых генов, у регуляторного белка мышц MyoD1 и у некоторых белков дрозофилы. Однако в этих случаях основной домен расположен по соседству с участком полипептидной цепи, образующим домен типа "спираль–поворот–спираль", в котором две амфипатические спирали (содержащие все заряженные аминокислотные остатки на одной стороне спирали) разделены неспирализованной петлей. Предполагают, что такой домен играет ту же роль в димеризации полипептидных цепей и формировании ДНК-связывающего центра, что и описанные выше регуляторные полипептиды, содержащие лейциновые домены.

Следует еще раз подчеркнуть, что структуры типа "лейциновая застёжка" и "спираль–поворот–спираль" необходимы для димеризации полипептидных цепей транскрипционных факторов, принадлежащих к данной группе, что сопровождается формированием основного ДНК-связывающего участка регуляторных белков. Белок протоонкогена *тус* содержит как лейциновый домен, так и домен типа "спираль–поворот–спираль" по соседству с основным

ДНК-связывающим доменом. В соответствии с этим все данное семейство регуляторных белков можно разделить на подсемейства, в которых полипептидные цепи факторов транскрипции содержат порознь или одновременно домены типа "лейциновая застежка" или "спираль–поворот–спираль".

Таковы особенности структуры некоторых наиболее хорошо изученных доменов факторов транскрипции, обеспечивающих специфичность взаимодействия последних с регуляторными последовательностями ДНК, что необходимо для их регуляторного воздействия на транскрипцию соответствующих генов. Количество известных доменов у факторов транскрипции быстро возрастает с развитием исследований в этой области. Так, новые ДНК-связывающие участки полипептидных цепей недавно обнаружены у фактора транскрипции AP2, факторов, обеспечивающих регуляторное действие сыворотки, белков STF/NTF, взаимодействующих с СААТ-боксом, факторов транскрипции дрожжей HAP-2 и HAP-3. В белках существует несколько структур, обеспечивающих специфическое распознавание определенных последовательностей ДНК, что позволяет осуществлять передачу регуляторных сигналов к определенным генам-мишеням и служит основой для обеспечения их дифференциальной экспрессии на уровне транскрипции. Данные о некоторых наиболее важных с функциональной точки зрения доменах факторов транскрипции суммированы в табл. I.14.

Несмотря на то что специфическое связывание с ДНК является обязательным этапом для регуляторного воздействия белковых факторов на транскрипцию, реализация активности факторов часто требует их специфического взаимодействия с другими регуляторными белками, а также с самой РНК-полимеразой II. Большинство таких воздействий приводит к активации транскрипции. Однако результатом некоторых из них может быть и репрессия синтеза мРНК. Ниже будут рассмотрены условия, необходимые для осуществления позитивного действия факторов транскрипции.

Для идентификации участков полипептидных цепей факторов транскрипции, непосредственно участвующих в активации синтеза РНК, чаще всего конструируют химерные рекомбинантные белки (некоторые из них кратко описаны выше), объединяющие в одной цепи ДНК-связывающий домен и

участки полипептидных цепей других факторов, и изучают влияние таких гибридных факторов на транскрипцию *in vitro*. В ходе исследований, основанных на подобных подходах, были идентифицированы функциональные домены факторов транскрипции, в том числе домены, участвующие в активации транскрипции, которые, как правило, отличаются от ДНК-связывающих участков их полипептидных цепей. Доменную структуру факторов транскрипции можно проиллюстрировать на примере семейства рецепторов стероидных/тиреоидных гормонов (рис. 1.26). В полипептидной цепи таких факторов четко видны три функциональных модуля: ДНК-

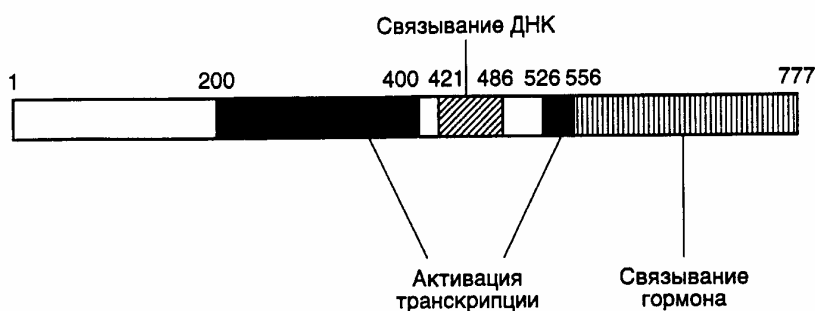


Рис. 1.26. Доменная структура полипептидной цепи рецептора глюкокортикоидов

Цифрами указаны номера аминокислотных остатков на границах функциональных доменов, участвующих в активации транскрипции, связывании ДНК и гормона

Таблица I.14

Функциональные домены факторов транскрипции

Домен	Функция	Факторы, содержащие домен	Примечание
Гомеобокс (гомеодомен)	Связывание ДНК	Гомеозисные гены <i>Drosophila</i> и других организмов	Взаимодействие с ДНК посредством домена "спираль–поворот–спираль"
POU	То же	Белки Oct-1, Oct-2, Pit-1 млекопитающих; ген <i>unc-86</i> нематод	Родственные гомеодоменам
Цистеин-гистидиновые "цинковые пальцы"	«	Белки TFIIIA, Kruppel, Sp1 и т.д.	Множественные копии домена "цинковые пальцы"
Цистеин-цистеиновые "цинковые пальцы"	«	Семейство рецепторов тиреоидных/стероидных гормонов	Одна пара "цинковых пальцев"; родственные мотивы имеются у E1A аденовирусов и GAL4 дрожжей
Основной элемент	«	Белки C/EBP, c-Fos, c-Jun, GCN4	Часто обнаруживают с доменами "цинковые пальцы", "спираль–поворот–спираль" или с ними вместе
"Лейциновая застежка"	Димеризация белков	Белки C/EBP, c-Fos, c-Jun, GC4N, c-Myc	Опосредуют димеризацию, необходимую для связывания ДНК

соседними доменами

Таблица I.13 (окончание)

Домен	Функция	Факторы, содержащие домен	Примечание
Спираль– поворот– спираль	То же	Белки с-Мус, MyoD дрозофилы; E12, E47 животных	То же
Амфипатическ ая кислая α - спираль	Активация генов	Дрожжевые белки GCN4, GAL4; рецепторы тиреоидных /стероидных гормонов	Прямое взаимодействие с белком TFIID
Область, обогащенная Gln	То же	Белок Sp1	Имеются гомологичные участки в белках Oct-1, Oct-2, Ap-2
Область, обогащенная Pro	«	Белок CTF/NF1	Имеются гомологичные участки в белках Oct-2, Ap-2, c-Jun

связывающий домен, домен, участвующий в связывании гормона, а также домены, необходимые для активации транскрипции. В отличие от этого белок VP16 вируса простого герпеса содержит активирующий участок полипептидной цепи, однако в нем отсутствует ДНК-связывающий домен. Белок VP16 образует специфический комплекс с клеточным ДНК-связывающим белком Oct-1, после чего происходит взаимодействие комплекса с ДНК и активация транскрипции. Таким образом, в этом случае ДНК-связывающий и активирующий домены комплексного фактора транскрипции локализованы в разных полипептидных цепях.

Сравнение структур активирующих доменов различных факторов транскрипции показало, что хотя подобные домены не обладают выраженной гомологией, все они обогащены кислыми аминокислотами. Эти аминокислотные остатки организованы таким образом, что образуют амфипатическую α -спираль, в которой все отрицательно заряженные остатки аминокислот расположены на поверхности спирали. Для подтверждения важной роли данной структуры в активации транскрипции пептид, образующий "кислую" (обогащенную остатками Asp, Glu) амфипатическую α -спираль, генно-инженерными методами соединяли с ДНК-связывающим доменом дрожжевого фактора транскрипции GAL4. Такой гибридный белок приобретал способность активировать транскрипцию. Однако этого не происходило, если с ДНК-связывающим доменом соединяли пептид, образующий неамфипатическую α -спираль, в которой те же самые аминокислотные остатки были расположены случайным образом (рис. 1.27). Хотя такие кислые активирующие домены обнаружены у многих активаторов транскрипции разных организмов (от дрожжей до человека), описаны и другие домены, выполняющие аналогичную функцию. Так, активирующий домен фактора Sp1 содержит участок, обогащенный остатками Gln. Однако в соответствующем домене факторов STF/NF1 преобладают остатки Pro. Такие Pro- и Gln-богатые участки обнаружены и у других факторов транскрипции. Следовательно, данная структура факторов транскрипции не является исключением.

В настоящее время становится все более очевидным, что различные активирующие домены осуществляют стимуляцию транскрипции с участием других белковых факторов, хотя тот же самый эффект может достигаться и

путем непосредственного их взаимодействия с РНК-полимеразой II. В частности, действие кислого активирующего домена опосредовано фактором TFIID, связывающим ТАТА-последовательность промоторных участков ДНК. Например, соединение дрожжевого фактора транскрипции GAL4 и фактора млекопитающих ATF со своими специфическими последовательностями в регулируемых промоторах меняют конформацию уже связанного с этим промотором фактора TFIID таким образом, что последний начинает контактировать не только с самим ТАТА-боксом, но и с последовательностью нуклеотидов вблизи точки инициации транскрипции. Как было упомянуто выше, такое изменение конформации фактора TFIID необходимо для вхождения в стабильный транскрипционный комплекс других факторов, в частности TFIIC и TFIIЕ, а также и самой РНК-полимеразы. Следовательно, специфические участки полипептидных цепей факторов транскрипции осуществляют свое активирующее действие и путем изменения конформации других факторов, связанных с промотором, что обеспечивает дальнейшую сборку стабильных транскрипционных комплексов.

Генетический контроль активности факторов транскрипции.

Первичная роль многих факторов транскрипции заключается в активации некоторых групп генов в определенных тканях в ответ на поступление специфических сигналов, например как следствие действия стероидных гормонов. Для достижения данной цели конкретные факторы транскрипции должны быть активны только в строго детерминированных тканях или в ответ на появление соответствующего сигнала. Включение фактора транскрипции в каскад этих реакций может быть достигнуто путем тканеспецифического синтеза соответствующего белка или регулируемой активации белка-предшественника в определенном месте и в заданное время (рис. 1.28).

Регуляция на уровне биосинтеза факторов транскрипции. Для многих регуляторных белков биосинтез ограничен клетками строго определенных типов. Например, октамерсвязывающий белок Oct-2, участвующий в активации экспрессии генов иммуноглобулинов в В-лимфоцитах, обнаруживают только в клетках, синтезирующих иммуноглобулины, но не в других клетках, например HeLa. Экспрессия рекомбинантного гена, кодирующего Oct-2, в клетках HeLa приводит к активации экспрессии генов иммуноглобулинов на уровне транскрипции. Таким образом, для осуществления тканеспецифической

регуляции экспрессии генов в этом примере необходим тканеспецифический синтез белка-активатора транскрипции (см. рис. 1.28,а).

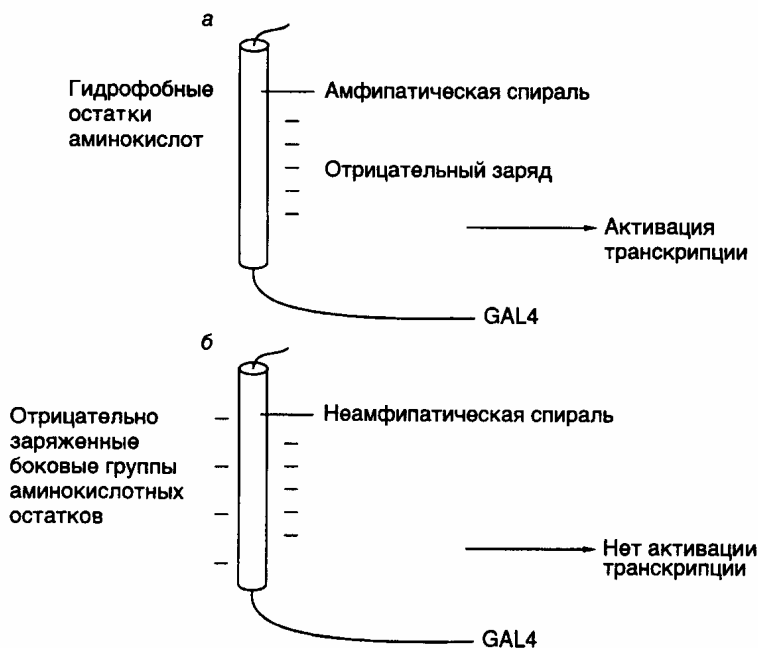


Рис. 1.27. Роль структуры типа амфипатической α -спирали в составе полипептидной цепи рекомбинантного фактора транскрипции GAL4
Транскрипция активируется за счет полипептидного домена, содержащего амфипатическую α -спираль (а), и не активируется, когда в той же самой аминокислотной последовательности отрицательно заряженные боковые группы распределены случайным образом (б)

Во многих случаях регулируемая экспрессия генов факторов транскрипции достигается путем соответствующего управления транскрипцией генов этих факторов. В частности, в клетках HeLa отсутствует не только белок Oct-2, но и его мРНК. Точно так же транскрипция гена фактора C/EBP происходит только в ядрах клеток печени. Очевидно, что такой способ контроля транскрипции у эукариот не решает полностью проблему регуляции экспрессии генов, поскольку предполагает необходимость наличия регуляторных генов, влияющих на транскрипцию генов факторов транскрипции, регулируемая экспрессия которых, в свою очередь, требует новых факторов, и так далее до бесконечности. В этой связи не является неожиданным, что регуляция экспрессии генов многих факторов транскрипции происходит на посттранскрипционном уровне. Например, возрастание уровня биосинтеза

дрожжевого фактора GCN4, активирующего гены биосинтеза аминокислот, является следствием ускорения трансляции его мРНК рибосомами в ответ на недостаток внутриклеточного содержания аминокислот.

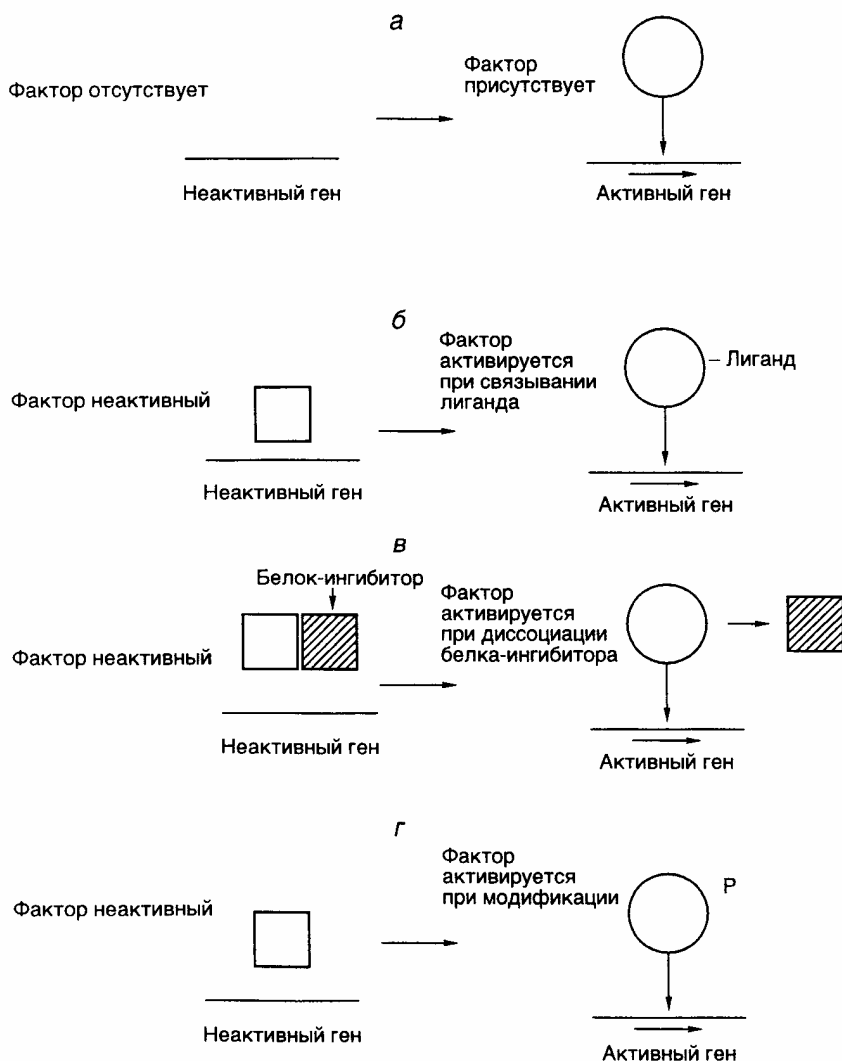


Рис. 1.28. Способы регуляции активности факторов транскрипции у эукариот (а–г)

Регуляция экспрессии генов факторов транскрипции может происходить и на уровне сплайсинга соответствующих РНК. В частности, существуют две формы рецептора тиреоидных гормонов, образующиеся в результате альтернативного сплайсинга. У одной из них отсутствует домен, связывающий гормон, и она способна распознавать те же последовательности ДНК, что и

гормонсвязывающий рецептор, однако не может активировать транскрипцию в присутствии гормона. Таким образом, эта форма действует как доминантный репрессор соответствующих генов. Подобный механизм описан также и для онкогена *v-erbA* вируса эритробластоэмии птиц, кодирующего укороченную форму рецептора тиреоидных гормонов, у которой отсутствует домен, связывающий гормон.

Регуляция активности факторов транскрипции. Белковые продукты генов многих специфически действующих факторов транскрипции часто присутствуют во всех тканях, однако специфический характер их воздействия достигается путем их посттрансляционной активации в строго определенном месте или же в ответ на соответствующий сигнал. Простым примером такого рода является активация дрожжевого фактора транскрипции ACE1, который стимулирует транскрипцию гена металлотионеина в присутствии ионов меди. В этом случае ионы Cu^{2+} , взаимодействуя с фактором, вызывают конформационные изменения в его полипептидной цепи, после чего фактор приобретает способность связываться с регуляторным участком гена металлотионеина и активировать его транскрипцию (см. рис. 1.28,б).

Аналогичную зависимость от активирующего лиганда демонстрируют молекулы факторов транскрипции, принадлежащие к семейству рецепторов стероидных/тиреоидных гормонов (см. рис. 1.26). Молекулы таких рецепторов для осуществления активирующего действия на гены-мишени должны вначале специфически связать эквимольные количества соответствующего гормона-эффектора. Как уже упоминалось, эти рецепторы обладают специальным С-концевым доменом, выполняющим данную функцию. Несмотря на то что *in vivo* такие рецепторы приобретают способность взаимодействовать с регуляторными последовательностями ДНК только в присутствии гормона, *in vitro* они связываются специфическими последовательностями ДНК как при наличии гормона, так и без него. Оказалось, что в клетках рецепторы находятся в комплексе с белком, предотвращающим их связывание соответствующими регуляторными последовательностями ДНК, и гормоны после взаимодействия с рецепторами в этих комплексах вызывают диссоциацию последних (см. рис. 1.28,в).

Следовательно, в рассмотренном выше случае активация факторов транскрипции происходит не в результате конформационного изменения их

пространственной структуры под действием лигандов, а путем лигандзависимого разрушения ингибирующего белок–белкового взаимодействия. Подобный механизм активации продемонстрирован для белкового фактора GAL4 в ответ на действие галактозы, а также белка NF κ B под воздействием форболовых эфиров в Т-лимфоцитах или клетках HeLa. Более того, аналогичный механизм обеспечивает активацию транскрипции генов в определенных тканях. Так, фактор транскрипции MyoD1 играет ключевую роль в активации экспрессии генов, происходящей в тканях мышц во время дифференцировки миобластов в мышечные волокна (миотубы). Такая активация наблюдается не из-за увеличения содержания MyoD1 в дифференцирующихся клетках, а как следствие уменьшения содержания белка-ингибитора Id, образующего комплекс с MyoD1 и препятствующего его взаимодействию с регуляторными последовательностями ДНК. Интересно, что ингибитор Id, как и сам MyoD1, содержит мотив "спираль–поворот–спираль", который опосредует димеризацию соответствующих белков. Однако в составе Id-белка отсутствует основной ДНК-связывающий домен, функционирование которого обсуждалось выше. Предполагается, что Id димеризуется с MyoD1, подавляя его способность взаимодействовать с ДНК, по аналогии с тем, как это происходит в опытах с укороченными белками MyoD1, не содержащими ДНК-связывающий домен.

Активация факторов транскрипции может осуществляться не только путем изменения белок–белковых взаимодействий, но и под действием ковалентных модификаций самих факторов в ответ на появление специфических сигналов (см. рис. 1.28,е). Примером может служить механизм активации фактора транскрипции CREB, который обеспечивает активацию некоторых клеточных генов в ответ на воздействие циклическим АМР. О подобных механизмах речь уже шла в разделе о вторичных мессенджерах. В этом случае сАМР стимулирует протеинкиназу А, которая, в свою очередь, фосфорилирует CREB, что сопровождается активацией домена, расположенного в полипептидной цепи фактора по соседству с сайтом фосфорилирования. Тот же механизм функционирует при активации фактора транскрипции генов теплового шока дрожжей в ответ на повышение температуры, а также при активации фактора NF κ B под действием форболовых эфиров. В последнем случае фосфорилирование белка-ингибитора,

находящегося в комплексе с белком NKкВ, приводит к диссоциации комплекса, что допускает последующее связывание NKкВ с регуляторными последовательностями ДНК, сопровождаемое активацией транскрипции соответствующих генов. Фосфорилирование не является единственной модификацией, приводящей к активации факторов транскрипции. Аналогичный эффект достигается и при гликозилировании некоторых белков.

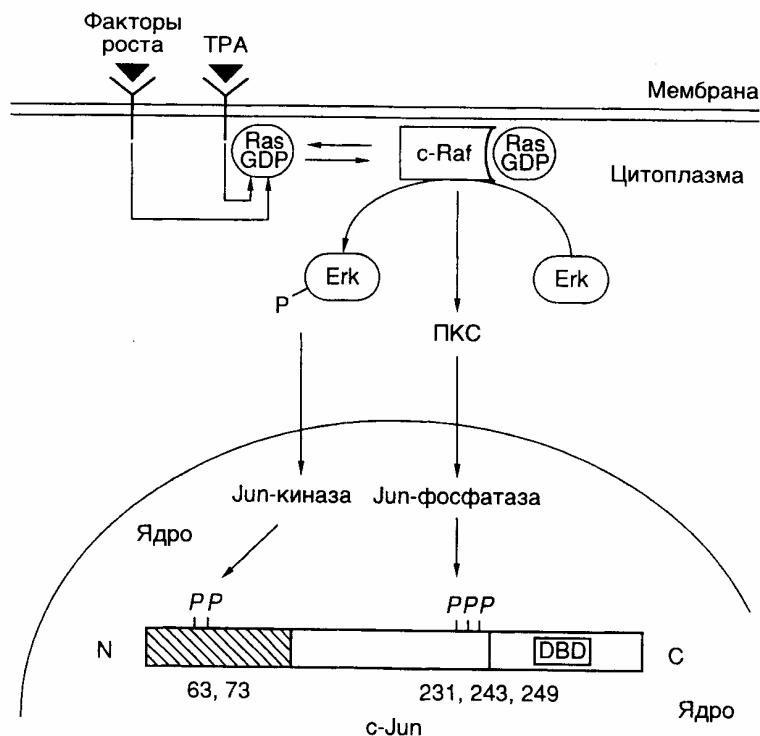


Рис. 1.29. Схема регуляции активности фактора транскрипции c-Jun

PKC – протеинкиназа C, DBD – ДНК-связывающий домен; 12-О-тетрадеканоилфорбол-13-ацетат. Цифрами обозначены номера остатков аминокислот в полипептидной цепи c-Jun, подвергающихся регуляторному фосфорилированию/дефосфорилированию. Заштрихован N-концевой домен, участвующий в активации транскрипции

В заключение рассмотрим несколько подробнее механизм регуляции активности факторов транскрипции класса 1.1 (факторы с доменами типа "лейциновая застежка"), которые, как уже упоминалось, относятся к Bzip-белкам, так как содержат основной домен (B – basic) и домен типа "лейциновая застежка". Данный пример поучителен потому, что приоткрывает завесу над некоторыми молекулярными механизмами канцерогенеза, поскольку многие белковые компоненты этой системы регуляции (Src, Ras, Raf, Jun, Sis, Fos)

являются продуктами экспрессии протоонкогенов.

Как уже упоминалось выше, семейство AP-1-подобных факторов транскрипции, предсуществующих в клетках в латентной форме, активируется классическим промотором химического канцерогенеза – 12-О-тетрадеканойлфорбол-13-ацетатом (ТРА), а также различными пептидными гормонами, факторами роста, цитокинами и нейромедиаторами. Одновременно происходит индукция транскрипции генов *c-fos*, уровень экспрессии которых в нестимулированных клетках низок. Латентная форма фактора *c-Jun* фосфорилирована вблизи С-концевого ДНК-связывающего домена и слабо фосфорилирована в N-концевом активирующем домене. Как следует из упрощенной схемы, представленной на рис. 1.29, эта последовательность реакций инициируется после взаимодействия с рецепторами на поверхности клетки соответствующих лигандов – цитокинов или факторов роста. Под действием протеинкиназ Src или PTK гуанозинтрифосфатаза Ras связывает молекулу GTP и переходит в активную конформацию, что позволяет эффекторному участку полипептидной цепи белка взаимодействовать с N-концевым доменом *c-Raf*. Процесс сопровождается переносом последнего к цитоплазматической мембране, его активацией и запуском каскада реакций, осуществляемых протеинкиназами, активируемыми митогенами (MAP), к которым относятся белки Erk-1 и Erk-2. Активность MAP-киназ зависит от фосфорилирования их полипептидных цепей по остаткам Thr и Ser киназой MAP-киназ (MAPKK), которая сама активизируется в результате фосфорилирования ее полипептидной цепи. Белок *c-Raf* в этом каскаде реакций выполняет роль такой MAPKK. Активированные Erk-1 и Erk-2 далее транслоцируются в ядро, где инициируют Jun-киназу, которая фосфорилирует фактор транскрипции *c-Jun* в N-концевом домене. Одновременно с участием протеинкиназы C происходит активация Jun-фосфатазы, которая дефосфорилирует полипептидную цепь *c-Jun* вблизи ее С-концевого домена. В результате реакций фосфорилирования и дефосфорилирования полипептидной цепи *c-Jun* этот фактор транскрипции приобретает способность взаимодействовать с регуляторными последовательностями генов, называемых ТРА-респонсивными элементами, и далее активирует транскрипцию соответствующих генов.

Рассмотрение механизмов позитивного контроля транскрипции

продолжим на примере механизмов функционирования регуляторных последовательностей нуклеотидов высших организмов – энхансеров – специфических регуляторных последовательностей, обеспечивающих высокий уровень транскрипции определенных генов после взаимодействия с регуляторными белками.

Энхансеры. Энхансерами называют определенный класс регуляторных последовательностей нуклеотидов, которые обладают рядом существенных особенностей, резко отличающих их от других регуляторных последовательностей эукариот, регулирующих транскрипцию. Энхансеры представляют собой протяженные последовательности нуклеотидов, которые содержат сайты связывания нескольких факторов транскрипции. Характерными свойствами энхансера являются его способность осуществлять регуляторное действие на промотор на больших расстояниях от него, достигающих 60 т.п.о. и более, независимость его активности от ориентации по отношению к промоторам, а также от расположения относительно регулируемого гена.

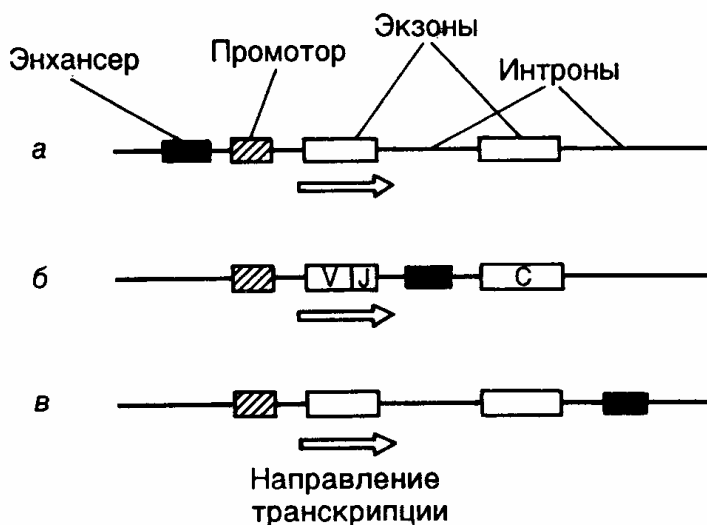


Рис. 1.30. Варианты взаимного расположения регуляторных и структурных частей генов эукариот

Регуляторные части генов представлены энхансером и промотором, структурные части – экзонами и интронами. а – ген альбумина, энхансер располагается перед промотором; б – ген иммуноглобулина, энхансер расположен в центре гена между последовательностями, кодирующими константную и переменную части белка; в – ген β -глобина, энхансер расположен вблизи кодирующей части гена

На рис. 1.30 приведены схемы строения нескольких эукариотических генов, которые отражают взаимное расположение их структурных и регуляторных частей. Если у гена альбумина (см. рис. 1.30,а) энхансер находится перед промотором, а вся регуляторная часть предшествует его структурной части, то в случае генов иммуноглобулинов регуляторные элементы локализованы в интронах самого гена (см. рис. 1.30,б). Энхансер может быть расположен и ниже гена на значительном от него расстоянии, как это имеет место у β -глобинового гена (см. рис. 1.30,в).

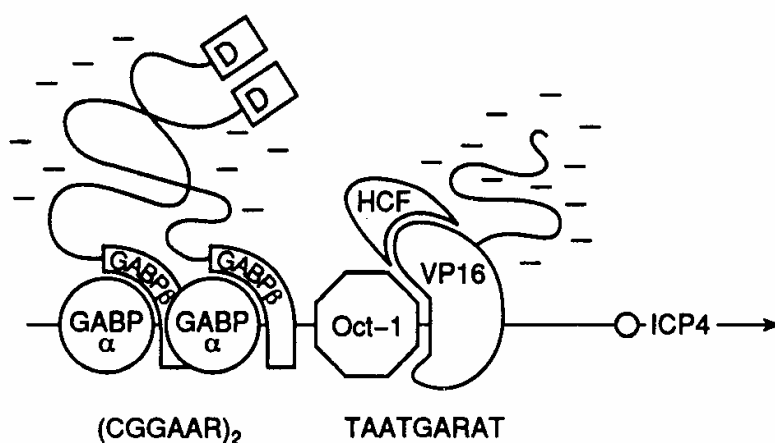


Рис. 1.31. Схема белок-белковых и белково-нуклеиновых взаимодействий на энхансерах непосредственно-ранних генов ВПГ
 D – домены димеризации соответствующих факторов,
 Отмечены отрицательно заряженные "кислые" участки их полипептидных цепей, непосредственно участвующие в активации транскрипции гена ICP4, направление транскрипции которого обозначено стрелкой

Исследование молекулярных взаимодействий факторов транскрипции с энхансерами привело к пониманию того, что такие регуляторные последовательности являются своеобразными матрицами для сборки сложных белковых комплексов, структура которых обеспечивает высокоспецифические белок-белковые связывания и передачу регуляторных сигналов РНК-полимеразе II, находящейся в составе инициационного комплекса. Как и в большинстве других случаев, основной прогресс в изучении молекулярных механизмов регуляции экспрессии генов с помощью энхансеров у эукариот был достигнут с использованием вирусов животных в качестве объектов исследования. У вируса простого герпеса (ВПГ) был обнаружен

транскрипционный фактор VP16, участвующий в активации его непосредственно-ранних генов и присутствующий в зрелом вирионе в количестве ~1000 копий. Белок VP16 взаимодействует с энхансерами, расположенными перед каждым из непосредственно-ранних генов вируса, что приводит к активации их транскрипции. Детальное изучение регуляторных участков этих генов показало, что каждый из энхансеров содержит одну или несколько копий *цис*-регуляторных последовательностей – нонануклеотидную последовательность 5'-TAATGARAT-3' (так называемая *tat-garat*) и прямые повторы гексануклеотидной последовательности 5'-CGGAAR-3' ("*cigar*"). В очищенном состоянии белок VP16 не взаимодействует ни с одной из них. Однако в присутствии белковых факторов экстрактов ядер клеток-хозяев VP16 входит в состав многокомпонентного белкового комплекса, формирующегося на последовательности "*tat-garat*" (рис. 1.31). Сходство последовательностей нуклеотидов мотива "*tat-garat*" и сайтов связывания Oct-белков, ранее обнаруженных в регуляторных последовательностях генов гистонов и иммуноглобулинов, заставило проверить эти белки на способность взаимодействовать с ВПГ-энхансерами. Было выяснено, что Oct-1 является вторым белковым компонентом, соединяющимся с "*tat-garat*"-последовательностью энхансера и облегчающим связывание с ним белка VP16.

Эффективное образование вышеупомянутого белкового комплекса требует участия других компонентов клетки-хозяина, поскольку очищенные белки Oct-1, VP16 и последовательность "*tat-garat*" формируют стабильный комплекс только в присутствии экстрактов ядер клеток HeLa. Белковый фактор, названный HCF, может непосредственно взаимодействовать с VP16 в отсутствие других компонентов комплекса. С помощью мутационного анализа было установлено, что 5'-концевой сегмент "*tat-garat*" (последовательность "*tat*") необходим для связывания белка Oct-1, тогда как ее 3'-концевая последовательность ("*garat*") требуется для образования полного белкового комплекса Oct-1–VP16–HCF. Таким образом, по крайней мере, четыре молекулярных поверхности (интерфейса) участвуют в формировании функционально активного комплекса на матричной поверхности мотива "*tat-garat*" энхансеров непосредственно-ранних генов ВПГ. Одна группа контактов

образуется между белками Oct-1 и VP16, вторая – между VP16 и HCF, третья – между Oct-1 и последовательностью "tat" и, наконец, четвертая – между VP16 и последовательностью "garat" (см. рис. 1.31).

Не менее сложная серия молекулярных интерфейсов участвует и в формировании функционального макромолекулярного комплекса на мотиве "cigar" энхансера непосредственно-ранних генов ВПГ. Белок, взаимодействующий с этим мотивом, выделен из экстрактов ядер печени крыс и получил название GABP (GA-binding protein). Он состоит из двух разных субъединиц – GABP α и GABP β . Белковые комплексы, похожие на GABP, выделены и из культивируемых клеток человека. Часть аминокислотной последовательности GABP α оказалась гомологичной участку ДНК-связывающего домена транскрипционных факторов, принадлежащих к семейству ETS. Все представители этого семейства обладают последовательностью из 85 аминокислотных остатков, которая необходима для их специфического взаимодействия с ДНК. Таким образом, ETS-домен белка GABP α образует первый молекулярный интерфейс, обеспечивающий связывание белка с мотивом "cigar" энхансера.

В смеси с GABP β α -субъединица образует белковый комплекс, который обладает бóльшим сродством к мотиву "cigar" энхансера, чем сама α -субъединица. Анализ особенностей взаимодействия GABP-комплекса с энхансером позволил обнаружить три высокоспецифичных молекулярных интерфейса. Это прежде всего домен полипептида GABP α , необходимый для его взаимодействия с белком GABP β и включающий ETS-домен, а также последовательность из 35 С-концевых аминокислотных остатков GABP α . У полипептидной цепи GABP β во взаимодействии с GABP α принимают участие четыре tandemных повтора, состоящие из 33 аминокислотных остатков каждый и расположенные в N-концевой части белка. Анализ аминокислотных последовательностей, гомологичных этому повтору у ряда других известных белков, позволил выявить высококонсервативную последовательность TPLN, из-за которой данный мотив и получил свое название. Дополнительные молекулярные взаимодействия между белковым комплексом и последовательностью нуклеотидов "cigar" осуществляются через повторы TPLN после того, как они образуют требуемый комплекс с GABP α . При этом

аминокислотные последовательности повторов непосредственно контактируют с ДНК. Наконец, в последнем из известных контактов, имеющих место после сборки белкового комплекса на последовательности нуклеотидов "cigar", участвует так называемый домен димеризации полипептидной цепи GABP β , расположенный вблизи ее С-конца. Очищенные субъединицы GABP β в растворе ассоциируют с образованием гомодимеров, которые в присутствии субъединиц GABP α формируют гетеродимерный комплекс $\alpha_2\beta_2$, хорошо приспособленный для взаимодействия с гексануклеотидным повтором "cigar" (CGGAAR)₂ энхансера.

Анализируя все вышеперечисленные компоненты энхансеров непосредственно-ранних генов ВПГ, можно сделать вывод, что в их активации принимают участие не менее пяти белков (VP16, Oct-1, HCF, GABP α и GABP β), которые участвуют, как минимум, в пяти белок-белковых контактах и шести специфических контактах белок–ДНК (см. рис. I.31). Прежде всего, обращает на себя внимание кооперативность взаимодействия компонентов, активирующих энхансеры. Ключевые белки этого комплекса – Oct-1, VP-16 и GABP α сами по себе слабо взаимодействуют со специфическими последовательностями ДНК, но обладают способностью собираться в стабильный функционально активный комплекс, по крайней мере, двумя путями: во-первых, объединяясь с кофакторами, которые самостоятельно не способны взаимодействовать с ДНК (например VP16 с HCF и GABP α с GABP β); во-вторых, контактируя с последовательностью нуклеотидов энхансера, содержащей функциональные участки "tat" и "garat", а также повтор "cigar" и образующей матрицу для геометрически правильной сборки всего активирующего комплекса.

При исследовании механизмов действия многочисленных факторов транскрипции удивляет то, что специфичность взаимодействия с ДНК многих белков этих подсемейств, по крайней мере, определенная *in vitro*, невысока. В очищенном состоянии факторы часто узнают одни и те же последовательности нуклеотидов. Каким же образом данные белки осуществляют специфические регуляторные воздействия на регулируемые гены? Как уже было видно при обсуждении функционирования энхансеров непосредственно-ранних генов ВПГ, такую специфичность обеспечивают другие, вспомогательные белки, взаимодействующие с основными регуляторными белками. В частности,

несмотря на сходство в специфичности распознавания последовательностей нуклеотидов, истинные регуляторные белки обладают бóльшим сродством к последовательностям-мишеням, еще более усиливающимся под влиянием белок-белковых взаимодействий. Недавно было показано, что белок GABP β не связывается с другими белками, обладающими ETS-доменами, и, следовательно, лишь один белок из этого семейства, а именно GABP α , образуя комплекс с GABP β , обеспечивает специфичность взаимодействия последнего с энхансером.

Рассмотренный нами пример регуляции экспрессии непосредственно-ранних генов ВПГ хорошо иллюстрирует те многочисленные белок-белковые и белково-нуклеиновые взаимодействия, которые требуются для функционирования сложной системы дифференциально экспрессирующихся генов высших организмов. Несмотря на то что с развитием молекулярной генетики число известных регуляторных механизмов будет увеличиваться, именно этот принцип, основанный на возможности специфических взаимодействий макромолекул друг с другом, по-видимому, останется доминирующим в регуляторных механизмах эукариот.

Коактиваторы транскрипции. Энхансеры обнаруживают одну из специфических черт регуляции транскрипции у эукариот: влияние на синтез РНК регуляторных последовательностей нуклеотидов, расположенных на больших расстояниях от инициаторной последовательности промоторов. Еще недавно совершенно непонятный механизм этого явления стал проясняться с открытием специфических регуляторных белков – коактиваторов транскрипции. Коактиваторы транскрипции были впервые обнаружены у дрожжей при исследовании одного из основных факторов транскрипции TFIID, который необходим для полноценной инициации транскрипции РНК-полимеразой II. Было установлено, что один из основных компонентов TFIID – TATA-связывающий белок ТВР способен обеспечивать базальную, но не индуцированную транскрипцию *in vitro*. На основании этого было высказано предположение, что он дополнительно должен содержать ТВР-ассоциированные факторы (TAF), которые отличаются от основных факторов и активаторов транскрипции. После клонирования генов TAF-факторов и очистки кодируемых этими генами белков выяснилось, что TAF-белки обеспечивают

физическое взаимодействие между активаторами транскрипции и основными факторами, поскольку они обладают способностью связывать белки обеих групп. Те же функции присущи и упомянутым выше SRB-белкам дрожжей. Таким образом, коактиваторы транскрипции являются белками-адаптерами, обеспечивающими перенос регуляторного сигнала от тканеспецифических белков-активаторов транскрипции к РНК-полимеразе II.

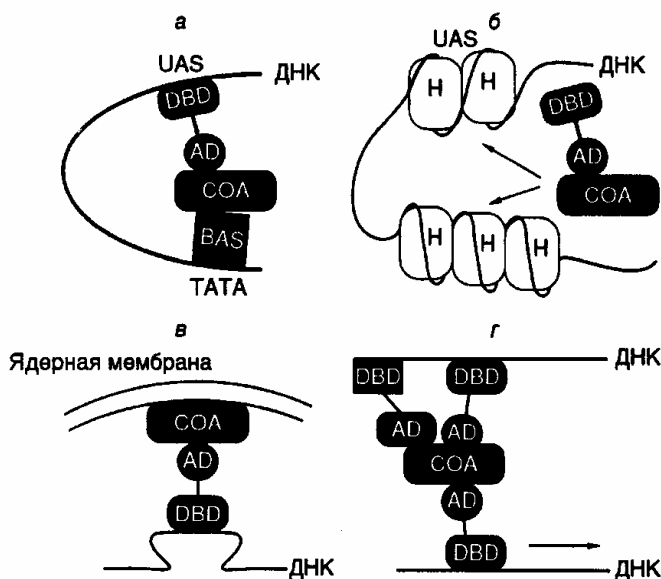


Рис. 1.32. Гипотетические модели механизма действия коактиваторов транскрипции

а – передача сигнала от активаторов транскрипции, взаимодействующих с дистальными регуляторными последовательностями нуклеотидов промоторов (UAS); **б** – разрушение нуклеосомной структуры в регуляторных последовательностях промоторов; **в** – поддержание определенной пространственной структуры транскрибируемых участков хроматина; **г** – обеспечение трансекции
AD и DBD – активирующий и ДНК-связывающий домены активатора транскрипции соответственно; COA – коактиватор транскрипции; BAS – основные факторы транскрипции; H – гистоны

На рис. 1.32 суммированы современные представления о механизмах действия коактиваторов транскрипции. Во всех четырех моделях подчеркивается адаптерная (медиаторная) функция этих белков, которые выступают в роли посредников между другими регуляторными белками. Прежде всего, акцентируется их роль в передаче сигнала от белков-активаторов транскрипции, реагирующих с дистальными регуляторными

последовательностями промоторов (UAS – upstream activation sites), к основным факторам транскрипции и РНК-полимеразе II (см. рис. 1.32,а). В этом случае коактиваторы транскрипции могут усиливать взаимодействие основных факторов транскрипции с инициаторными последовательностями промотора. Имеются экспериментальные данные, указывающие на то, что другая функциональная роль коактиваторов транскрипции заключается в дерепрессии промоторов, доступ к которым для основных факторов транскрипции блокирован гистонами, входящими в состав нуклеосом (см. рис. 1.32,б). В этом случае коактиваторы транскрипции могут участвовать в разрушении структуры хроматина в окрестностях промоторов. После удаления гистонов или изменения структуры нуклеосом регуляторные последовательности становятся доступными для основных факторов транскрипции и РНК-полимеразы II. Коактиваторы транскрипции могут играть определенную роль и в создании топологии транскрибируемых участков хроматина в интерфазных ядрах (см. рис. 1.32,в). В этом случае коактиваторы удерживают активно транскрибируемые участки хроматина в тех микрокомпартаментах интерфазных ядер, которые обеспечивают эффективную экспрессию генов. Такая роль коактиваторов отчетливо проявляется в функционировании инсуляторов и сайленсеров (см. ниже, а также раздел 3.2.3). И, наконец, последняя модель (см. рис. 1.32,г) объясняет механизм явлений, аналогичных процессу трансекции, впервые описанному у дрозофилы. Сущность этого явления заключается в том, что взаимодействие двух различных мутантных аллелей одного гена, присутствующих на разных гомологичных хромосомах, приводит к функциональной компенсации мутаций и усилению экспрессии соответствующего гена, т.е. трансекция является одной из разновидностей *генетической комплементации*. Трансекция происходит, когда, например энхансер гена, структурная часть которого инактивирована мутацией, активирует транскрипцию другого аллеля этого гена с поврежденным мутацией энхансером, находящегося на гомологичной хромосоме. В таком случае трансхромосомная активация транскрипции обеспечивается действием белков-коактиваторов транскрипции, с участием которых происходит передача регуляторного сигнала от активатора транскрипции, ассоциированного с энхансером одной хромосомы, к основным факторам транскрипции гомологичного аллеля этого гена другой хромосомы. Белки-коактиваторы

обеспечивают и синапсис (конъюгацию, объединение) гомологичных хромосом на протяженном участке.

Наличие у эукариот регуляторных белков-коактиваторов транскрипции объясняет одну из сторон механизма действия энхансеров, расположенных дистально по отношению к точке инициации транскрипции. В соответствии с этой концепцией дистальные и проксимальные регуляторные элементы промотора, сближенные в результате образования петли, разделяющей их ДНК, удерживаются вместе коактиваторами транскрипции. Однако остается необъясненной еще одна сторона механизма действия энхансеров. Действительно, энхансеры как регуляторные последовательности нуклеотидов обладают рядом уникальных свойств. Их стимулирующее действие на промоторы регулируемых генов не зависит от ориентации, а также положения по отношению к гену и проявляется даже в том случае, если энхансеры удалены от генов-мишеней на значительное (несколько десятков тысяч пар оснований) расстояние. Последнее обстоятельство кажется наиболее загадочным, так как непонятно, каким образом удается избежать активации промоторов гетерологичных генов, которые могут быть расположены в пределах досягаемости действия этих энхансеров. Как оказалось, специфичность действия энхансеров обеспечивается еще одним весьма своеобразным механизмом регуляции экспрессии генов у высших организмов, в котором участвуют специфические регуляторные последовательности нуклеотидов – так называемые пограничные последовательности, или инсуляторы (раздел 3.2.4).

3.2.3. Механизмы негативной регуляции транскрипции

Позитивный контроль транскрипции у эукариот, в котором участвуют многочисленные активаторы транскрипции, играет ключевую роль в регуляции экспрессии их генов на уровне транскрипции. Однако негативная регуляция активности генов у эукариот является столь же жизненно важным регуляторным механизмом, как и у бактерий. Подавление транскрипции необходимо при установлении разделенных во времени и пространстве паттернов транскрипции в клетках различных тканей в онтогенезе, а также при изменении уровней синтеза РНК в ответ на регуляторные изменения в микроокружении клеток. Механизмы негативной регуляции экспрессии генов на уровне транскрипции у

эукариот разнообразны. Некоторые наиболее важные из них будут рассмотрены ниже.

Задержка транспорта факторов транскрипции из цитоплазмы в ядра. Подавление определенных генов может достигаться за счет задержки в цитоплазме соответствующих факторов транскрипции. В хорошо изученном случае белки-активаторы семейства Rel задерживаются в цитоплазме и не переносятся в ядра вследствие их взаимодействия с белковыми факторами I κ B. Диссоциация комплексов Rel-I κ B, сопровождаемая транспортом Rel-факторов в ядра, контролируется фосфорилированием белков I κ B. Rel-семейство факторов транскрипции включает в себя продукт онкогена v-Rel и клеточный гомолог c-Rel, а также морфоген дрозофилы dorsal и ядерный фактор NF κ B, взаимодействующий с энхансером B гена иммуноглобулина κ . Последний представляет собой гетеродимер, состоящий из субъединиц с молекулярными массами 50 и 65 кДа, и обеспечивает тканеспецифическую экспрессию генов в зрелых B-лимфоцитах. Фактор NF κ B участвует в активации транскрипции некоторых генов в ответ на внешние сигналы (например под действием цитокинов) в нелимфоидных клетках.

Все I κ B-белки содержат один и тот же участок, гомологичный структурному белку анкирину и необходимый для их взаимодействия с Rel-белками. В результате такого контакта маскируется сигнальная последовательность Rel-белков, которая важна для их транспорта в ядра и активации транскрипции соответствующих генов.

Активация факторов транскрипции из пула неактивных цитоплазматических комплексов обеспечивает быструю индукцию транскрипции в ответ на внешние регуляторные сигналы. Более того, эта регуляторная система предусматривает в случае необходимости и инактивацию Rel-факторов в ответ на соответствующие сигналы через изменение уровня фосфорилирования I κ B-белков. Поскольку имеется целое семейство I κ B-белков, обладающих разной специфичностью в отношении Rel-факторов, и различные белки семейства реагируют на разные внешние сигналы, такая система негативной регуляции является очень гибкой. Другим примером цитоплазматической задержки факторов транскрипции служат рецепторы глюкокортикоидов, которые в отсутствие соответствующего лиганда

(гормона) находятся в комплексе с белком теплового шока Hsp90 и не переносятся в ядра. В то же время этот белок способствует связыванию лиганда с неактивным рецептором глюкокортикоидов, сопровождаемому распадом комплекса, т.е. его функция не ограничивается ингибирующим действием.

Предотвращение взаимодействия факторов транскрипции с регуляторными последовательностями на ДНК. Внутрядерное предотвращение связывания активаторов транскрипции с соответствующими регуляторными последовательностями на ДНК является еще одним широко распространенным механизмом негативной регуляции транскрипции у эукариот. В некоторых случаях негативно действующий фактор связывается с регуляторной последовательностью нуклеотидов, которая располагается по соседству с позитивным регуляторным элементом или перекрывается с ним. Это создает стерические препятствия для связывания последним активатора транскрипции, что предотвращает индукцию синтеза РНК. Функционирование такого в основном пассивного механизма зависит от взаимного расположения негативных и позитивных регуляторных последовательностей в промоторе. В данном случае для проявления активности негативного регулятора транскрипции необходимо его связывание с соответствующей последовательностью ДНК, и его действие не распространяется на другие регуляторные последовательности.

Наиболее простой случай репрессии синтеза РНК описан для промотора гена β -интерферона, где для активации гена необходимо связывание двух позитивно действующих факторов. Другой белковый фактор, ассоциируя с таким участком ДНК, предотвращает взаимодействие с ним этих двух позитивных факторов и подавляет транскрипцию. В ответ на вирусную инфекцию негативно действующий фактор инактивируется, и происходит активация гена на уровне транскрипции под действием двух позитивных факторов. В другом механизме, известном под названием *сквелчинга* (squelching – подавление), сильный *транс*-активирующий белок (В) связывает все молекулы обычного фактора транскрипции (А) в растворе. Это предотвращает взаимодействие фактора транскрипции А с другими генами, содержащими сайты связывания только для фактора А, но не для *транс*-активирующего фактора В. Такой теоретически возможный механизм пока

экспериментально продемонстрирован только *in vitro*. Многие факторы транскрипции, принадлежащие к одному семейству, обладают похожей или идентичной специфичностью в отношении регуляторных последовательностей ДНК. Поэтому они конкурируют друг с другом за места связывания на ДНК, если присутствуют в одной и той же клетке. Типичным примером являются белки дрозофилы, содержащие гомеодомены, специфичности которых в значительной мере перекрываются. В частности, фактор, связывающий cAMP-респонсивный элемент (CREB), и белок-активатор 1 (AP-1) гомологичны в структурном отношении и узнают одни и те же последовательности ДНК. CREB способен взаимодействовать с каноническими сайтами AP-1, однако не может активировать транскрипцию с этих сайтов. Одновременное присутствие в ядре этих факторов подавляет AP-1-индуцированную транскрипцию. Уровень антагонизма между данными двумя факторами регулируется cAMP-зависимым фосфорилированием CREB.

Рецептор тиреоидных гормонов, активируемый лигандом (T_3R), является стимулятором транскрипции, который узнает палиндромную последовательность AGGTCA–TGACCT. Этот рецептор узнает также респонсивный элемент рецептора эстрогенов AGGTCA–NNN–TGACCT, однако не способен активировать транскрипцию на данном регуляторном сайте. Тем не менее, он конкурирует с рецептором эстрогенов за регуляторный сайт и подавляет синтез РНК, индуцируемый эстрогенами. Эстрогеновый респонсивный элемент окситоцинового промотора негативно регулируется рецептором ретиноевой кислоты по тому же механизму.

Образование гетеродимеров между субъединицами активаторов транскрипции. Многие факторы транскрипции взаимодействуют с ДНК в виде димеров. В димеризации участвуют специализированные домены этих белков, такие как "лейциновая застежка", POU-домен или мотив "спираль–поворот–спираль". Имеется определенная неразборчивость в образовании гетеродимеров внутри родственных семейств факторов транскрипции. Например, прототип фактора AP-1 представляет собой гетеродимер продуктов протоонкогенов c-Fos и c-Jun, которые взаимодействуют друг с другом доменами типа "лейциновая застежка". В клетках обнаружены, по крайней мере, четыре Fos- и три Jun-подобных белка, которые образуют различные гомо- и гетеродимерные комбинации внутри семейств и между ними. Разные

комбинации этих факторов обеспечивают широкое разнообразие регуляторных взаимодействий между гетеродимерами и регуляторными последовательностями нуклеотидов.

Регуляторный белок Id у млекопитающих и его гомолог у дрозофилы – продукт гена *ets* – являются антагонистами активаторов транскрипции, обладающих доменами типа "спираль–поворот–спираль". Эти белки взаимодействуют с факторами транскрипции, однако сами не имеют ДНК-связывающих доменов и, следовательно, подавляют активность своих партнеров по димеризации. Подобные регуляторные воздействия играют важную роль в онтогенезе данных организмов. В случае других негативно действующих факторов образующийся димер взаимодействует с регуляторными последовательностями промоторов. Однако он утрачивает способность реагировать на внешние активирующие сигналы. В двух последних примерах действие негативного белка-регулятора не ограничивается пассивной конкуренцией с активатором транскрипции за место посадки на ДНК, но способствует переводу белка-активатора в неактивную форму в виде гетеродимерного комплекса.

Интересным примером негативной регуляции транскрипции у животных является направленное изменение связывания продукта протоонкогена *c-Myc* с регуляторными последовательностями. Для активирующего взаимодействия с промотором и проявления своих онкогенных свойств белок-активатор транскрипции *c-Myc* должен образовать гетеродимерный комплекс с продуктом гена *max* (*myc auxiliary factor*). Однако, находясь в стехиометрическом избытке, белок *Max* становится ингибитором транскрипции соответствующих генов, образуя гомодимеры и связываясь с теми же регуляторными последовательностями ДНК, но не активируя транскрипцию. Этот пример иллюстрирует возможность сильного влияния на уровни транскрипции путем небольших изменений в относительном внутриклеточном содержании димеризующихся с образованием гетеродимеров белков-активаторов. Активирующее действие *Myc*-белка на транскрипцию регулируется взаимодействием *Max*-фактора с другими белками-антагонистами: *Myc*, *Mad* и *Mxi1*. Эти белки образуют с белком *Max* гетеродимеры, которые могут связываться регуляторными последовательностями ДНК, но не способны активировать транскрипцию. Такие белки конкурируют с белком *Myc* за его

активирующий кофактор Max и в составе гетеродимеров блокируют регуляторные последовательности. Известны и другие примеры негативной регуляции транскрипции у эукариот по данному механизму.

Многие негативно действующие факторы транскрипции являются продуктами генов, которые одновременно кодируют и активаторы транскрипции. Такие факторы-репрессоры синтезируются в результате альтернативного сплайсинга предшественника их мРНК или при альтернативном использовании иницирующих кодонов во время трансляции. Образование позитивно или негативно действующего фактора транскрипции во время экспрессии одного и того же гена регулируется в онтогенезе, в процессе дифференцировки тканей или в ответ на определенные внешние сигналы в конкретной группе клеток.

Взаимодействие между факторами, принадлежащими к разным классам. Некоторые факторы транскрипции в процессе эволюционных преобразований приобрели способность к белок-белковым взаимодействиям не только в пределах своего семейства, но и к перекрестным реакциям с членами других семейств. Подавлением транскрипции может сопровождаться образование специфических комплексов между факторами транскрипции, принадлежащими к разным классам, например между c-Jun и миогенным фактором D (MyoD) или между c-Jun и рецептором глюкокортикоидов. При этом в результате образования гетерогенных комплексов меняются как ДНК-связывающая активность факторов, так и их способность активировать транскрипцию без изменения ДНК-связывающих свойств белков-мишеней.

Белки, как обладающие, так и не обладающие ДНК-связывающей активностью, могут негативно регулировать транскрипцию путем маскировки активирующих поверхностей полипептидных цепей активаторов транскрипции, с которыми они взаимодействуют. Так, белок-регулятор метаболизма галактозы у дрожжей GAL80 самостоятельно не связывается с ДНК, но способен маскировать активирующий домен позитивного фактора транскрипции GAL4. Точно так же продукт ретинобластомного гена Rb не взаимодействует с ДНК непосредственно, но модулирует активность ДНК-связывающих факторов транскрипции. Поскольку каскад реакций, связанных с действием белка Rb, может оказывать негативное влияние на здоровье человека, вызывая в детском возрасте развитие онкологического заболевания – ретинобластомы,

рассмотрим механизмы этих реакций более подробно.

Белок-супрессор опухолевого роста Rb контактирует со многими клеточными белками, включая факторы транскрипции Sp1, ATF2, c-Мус и Elf-1. Наиболее хорошо изучен механизм его взаимодействия с фактором транскрипции E2F, вначале считавшимся регулятором транскрипции гена E2 аденовирусов. Сайты связывания E2F обнаружены на промоторах различных генов, регулируемых на протяжении клеточного цикла. К ним относятся, в частности гены *cdc2*, тимидинкиназы, ДНК-полимеразы α , *c-myb*, *c-myc* и дигидрофолатредуктазы.

Активатор транскрипции E2F образует многокомпонентные ДНК-связывающие комплексы, состав которых меняется во время клеточного цикла. Именно в фазе G1 в комплекс входит белок Rb, уровень фосфорилирования которого, меняющийся на протяжении клеточного цикла, в этой фазе минимален. Такое взаимодействие индуцирует негативную активность белка-супрессора опухолей Rb, поскольку именно в фазе G1 Rb вызывает задержку клеточного цикла. Кроме того, онкобелки онкогенных вирусов (Т-антиген вируса SV40, белок E7 вируса папилломы и белок E1a аденовирусов) во время вирусной инфекции взаимодействуют исключительно с не полностью фосфорилированным белком Rb. Это указывает на то, что изменение свойств белка Rb может быть одним из необходимых условий для опухолевого перерождения зараженных клеток. Онкобелки нарушают взаимодействие Rb с фактором транскрипции E2F, тем самым устраняя супрессорные свойства Rb (как негативно действующего белка-регулятора), которые индуцируются в нем в результате такого взаимодействия. При этом мутантные онкобелки, не способные нарушать связь между Rb и E2F, онкогенной активностью не обладают.

Фактор E2F является эффективным специфическим активатором транскрипции. Белок Rb дикого типа ингибирует активацию транскрипции фактором E2F, тогда как мутантный нефосфорилированный Rb представляет собой исключительно сильный ингибитор фактора. При этом Rb не только оказывает влияние на активацию транскрипции фактором E2F, но и превращает комплекс E2F–Rb в специфический репрессор синтеза РНК.

Таким образом, в соответствии с рассмотренной выше упрощенной моделью, Rb-белок в покоящихся клетках находится в комплексе с активатором

транскрипции E2F, подавляя экспрессию генов-мишеней этого фактора, которая необходима для вступления клеток в митоз. Сигналы, запускающие пролиферацию клеток, приводят к фосфорилированию Rb, диссоциации комплекса E2F-Rb и активации E2F как позитивного фактора транскрипции. Нормальная регуляция клеточного цикла может быть нарушена, по крайней мере, двумя онкогенными воздействиями: онкобелками, вытесняющими Rb из комплекса, и мутациями, нарушающими процесс взаимодействия Rb с фактором E2F.

Сайленсеры. Все рассмотренные выше способы негативной регуляции транскрипции по сути являются пассивными, так как лишь механически вмешиваются в разные этапы ее активации, нарушая их правильный ход. В то же время ингибирование транскрипции с использованием особых регуляторных элементов, называемых сайленсерами, – активный процесс. В этом случае происходит прямое подавление инициации транскрипции путем разрушения транскрипционного комплекса на промоторе или посредством его инактивации иным способом.

Первый из описанных в 1986 г. сайленсеров обладал классическими энхансероподобными свойствами, действуя на промоторы, расположенные в *цис*-положении (на той же молекуле ДНК) на большом расстоянии. При этом активность сайленсера, подобно энхансеру, не зависела от его ориентации по отношению к регулируемому промотору. Активность других сайленсеров в разной степени зависит от положения их по отношению к регулируемому промотору и ориентации относительно него, а также прямо пропорциональна числу их копий. Кроме того, регуляторные белки, связывающиеся с сайленсерами, по аналогии с белками энхансеров, помимо ДНК-связывающих доменов содержат аминокислотные последовательности, обеспечивающие белок-белковые взаимодействия, которые необходимы для осуществления негативной регуляции транскрипции. Исследование структуры этих доменов выявило их большое разнообразие, что позволяет думать о высокой функциональной значимости негативной регуляции транскрипции, обеспечиваемой сайленсерами.

3.2.4. Структура хроматина как специфический регулятор экспрессии генов

Регуляция тканеспецифической экспрессии генов с использованием

энхансеров и сайленсеров, а также некоторых других негативных регуляторных процессов часто происходит по механизмам, принципиально отличающимся от ранее рассмотренных. Речь идет о регуляторных эффектах, реализующихся через изменение структуры хроматина.

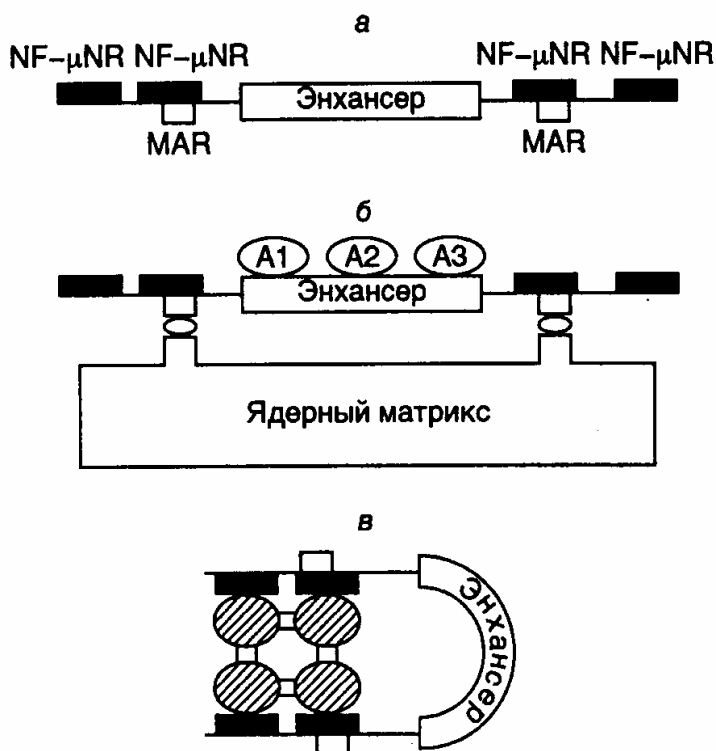


Рис. 1.33. Модель негативной регуляции активности интронного энхансера гена тяжелой цепи иммуноглобулина мышей

а – структура фрагмента ДНК в окрестностях энхансера. Обозначены сайты связывания фактора NF-μNR, а также MAR-последовательности; б – активное состояние энхансера, при котором MAR-последовательности ассоциированы с ядерным матриксом, а сам энхансер взаимодействует с активаторами А; в – неактивное состояние энхансера, при котором тетрамер фактора NF-μNR предотвращает взаимодействие MAR-последовательностей с ядерным матриксом и изменяет пространственную структуру энхансера

Интронный энхансер гена тяжелой цепи иммуноглобулина содержит много сайтов связывания для различных тканеспецифических, а также других широко распространенных факторов транскрипции. Максимальная активность энхансера наблюдается в зрелых В-лимфоцитах, где ген претерпевает ряд соматических перестроек. Активность этого энхансера обнаруживается и в

некоторых нелимфоидных тканях. Однако в печени или фетальных фибробластах, а также в нескольких других тканях его активность полностью подавлена. Исследования с использованием сайт-специфического мутагенеза показали, что такой энхансер фланкирован негативными регуляторными последовательностями нуклеотидов, которые ингибируют его активность в незрелых В-клетках или клетках других типов, но не обладают аналогичными свойствами в зрелых В-лимфоцитах (рис. I.33,а). Оба этих регуляторных элемента требуются для проявления ингибирующей активности, для которой важно также их положение относительно энхансера. В регуляторных элементах обнаружены сайты связывания транскрипционного фактора NF- μ NR, содержание которого особенно велико в клетках, где активность энхансера подавлена. Поскольку фактор NF- μ NR способен образовывать тетрамеры, предполагается, что прямое взаимодействие между этими молекулами, ассоциированными с сайтами связывания, фланкирующими энхансер, может приводить к структурной перестройке сегмента ДНК, заключенного между такими сайтами и содержащего энхансер (см. рис. I.33,б,в). Кроме того, регуляторные последовательности, связывающие NF- μ NR, содержат MAR-сайты, взаимодействующие с ядерным матриксом, и последовательности, с которыми контактирует ДНК-топоизомераза II, что еще раз указывает на необходимость структурных перестроек ДНК в процессе изменения ее регуляторных функций.

В рассматриваемом случае механизм негативной регуляции активности энхансера можно представить в следующем виде. В клетках, где отсутствует фактор транскрипции NF- μ NR (зрелые В-лимфоциты), последовательности MAR, которые фланкируют энхансер, ассоциированы с ядерным матриксом, что сопровождается формированием пространственной структуры энхансера, открытой для взаимодействия с активаторами транскрипции (см. рис. I.33,б). Это приводит к эффективной транскрипции всего гена иммуноглобулина. В клетках же, где фактор NF- μ NR присутствует в больших количествах, фрагмент ДНК с энхансером не ассоциирован с ядерным матриксом, так как MAR-сайты блокированы фактором (см. рис. I.33,в). Энхансер приобретает закрытую для позитивных регуляторных факторов конформацию, что сопровождается подавлением транскрипции регулируемого им гена. Рассмотренный пример

иллюстрирует один из частных случаев изменения пространственной структуры ДНК и хроматина, а также уровня транскрибируемости ДНК под действием коротких последовательностей нуклеотидов, фланкирующих регулируемый генетический локус. Подобный принцип широко используется для регуляции экспрессии генов у эукариот. Другие примеры этого впечатляющего явления в обобщенном виде будут рассмотрены ниже.

Эффект положения и концепция пограничных последовательностей: инсуляторы. Давно известно, что перенос гена или группы генов в гетерохроматиновые (неактивные в отношении транскрипции) участки хромосом часто сопровождается ослаблением или прекращением его экспрессии (так называемый эффект положения), и, наоборот, некоторые гены после переноса сохраняют свою активность и в гетерохроматиновом окружении. Подавление экспрессии таких транслоцированных генов может быть полным при стабильном эффекте положения и варьировать в зависимости от типа соматических клеток, в которых находится хромосома с транслокациями. Поскольку в последнем случае в организме образуются клоны соматических клеток, с разной эффективностью экспрессирующих транслоцированные гены (вплоть до полного подавления их транскрипции), такой эффект положения называют эффектом положения мозаичного типа (position effect variegation – PEV). Дальнейшее подтверждение существования эффекта положения получено с развитием методов трансгеноза – введения чужеродных рекомбинантных генов в геном клеток зародышевой линии высших организмов.

Исследование этого явления привело к открытию так называемых пограничных последовательностей нуклеотидов (boundaries), фланкирующих функционально активные домены хроматина. Оказалось, что существуют определенные последовательности нуклеотидов длиной в несколько сотен пар оснований, которые обладают способностью подавлять позитивное и негативное влияние эухроматина и гетерохроматина на экспрессию трансгенов, интегрированных в этот хроматин и фланкированных указанными последовательностями в новом сайте интеграции. Фактически такие участки ДНК как бы изолируют ген, находящийся между ними, способствуя сохранению его обычной пространственной структуры, которая может отличаться от структуры окружающего хроматина. Эти последовательности известны кроме

того под названием инсуляторов (англ. *insulate* – изолировать), а также как регуляторные области локусов (*locus control regions* – LCR). К таким пограничным последовательностям относятся, например А-элементы, фланкирующие ген лизоцима цыплят, scs-элементы (*specialized chromatin structure elements*), окружающие ген *hsp70* *Drosophila melanogaster*, а также последовательности нуклеотидов, разделяющие регуляторные элементы *iab* комплекса *Bithorax* того же объекта. Введение одного из таких элементов между энхансером и промотором регулируемого гена приводит к функциональной изоляции энхансера и подавлению экспрессии гена, а фланкирование гена пограничными последовательностями предохраняет его от инактивирующего действия окружающего конденсированного гетерохроматина, т.е. снимает эффект положения. При подавлении активности энхансеров инсуляторами ярко проявляется еще одно их свойство – полярность действия. Инсуляторы односторонне выключают энхансеры, расположенные дистально (на значительном расстоянии) по отношению к регулируемому промотору, но не рядом с ним (подробнее см. ниже). В дополнение к этим функциональным свойствам инсуляторов показано, что они могут разделять два участка хроматина, резко различающиеся по пространственной структуре. В этом случае по одной стороне от пограничной последовательности располагается сильно компактизованный хроматин, ДНК которого недоступна действию нуклеаз, а по другой – хроматин в открытой конформации, характерной для компетентных в отношении транскрипции генов.

Способность последовательностей нуклеотидов ДНК выполнять функции инсуляторов обычно определяют в функциональных генетических тестах по их способности обеспечивать экспрессию генов, искусственно интегрированных в хромосому (трансгенов), независимо от локализации сайта интеграции. В качестве гена-репортера при определении активности инсуляторов часто используют ген *white* дрозофилы, поскольку изменение уровня его экспрессии сопровождается легко измеряемым изменением окраски глаз у трансгенных мух, которая может меняться от красного, свойственного организмам дикого типа, до желтого и, наконец, белого, встречающегося у так называемых нуль-мутантов, для которых характерно полное подавление экспрессии мутантного гена. В такого рода исследованиях была обнаружена высокая эволюционная консервативность инсуляторов и регуляторных белков, взаимодействующих с

ними. В частности, LCR-последовательности β -глобинового гена цыплят эффективно функционируют в этой системе в качестве инсуляторов. Одной из наиболее хорошо изученных последовательностей нуклеотидов, обладающей свойствами инсулятора, является транскрибируемый нетранслируемый участок ДНК ретротранспозона *gypsy* у дрозофилы, содержащий сайты связывания белка-супрессора *Hairy-wing* (*su(Hw)*), свойства которого будут подробнее рассмотрены ниже.

До настоящего времени окончательно не разрешен вопрос о способности MAR/SAR-последовательностей, разделяющих функциональные домены хроматина, выполнять функции пограничных последовательностей. Хотя показано, что некоторые SAR-последовательности не способны подавлять эффект положения, для других аналогичных последовательностей такая способность была продемонстрирована. Например, свойствами инсуляторов обладают соответствующие последовательности уже упомянутого выше гена тяжелой цепи иммуноглобулина, а также генов лизоцима цыплят, β -интерферона и аполипопротеина В человека. В последнем случае сайты прикрепления ДНК к ядерным мембранам расположены за 5 т.п.о перед точкой инициации транскрипции и на 43 т.п.о. ниже этой точки, на границах между конденсированной и открытой формами хроматина. Последовательности, располагающиеся в этих участках хроматина, обладают повышенной чувствительностью к ДНКазе I и обеспечивают правильную экспрессию интегрированного в хромосому гена-репортера в клетках-трансфектантах гепатомы, вызывая 200-кратную стимуляцию его транскрипции, не зависящую от локализации места интеграции. Все это указывает на то, что данные 3'- и 5'-концевые MAR *APOB*-локуса человека обладают всеми свойствами пограничных последовательностей. Функционирование SAR-последовательностей в качестве инсуляторов в ряде случаев не является конститутивным, но зависит от типа клеток и стадии развития организма. Например, SAR-последовательности гена β -интерферона человека, фланкирующие мышинный ген теплового шока HSP70.1, функционируют как инсуляторы в клетках эмбрионов мышей на стадии предимплантации, но не в дифференцирующихся тканях новорожденных и взрослых трансгенных мышей. На такого рода наблюдениях основано заключение, что топологические домены

хромосом, выделяемые MAR/SAR-последовательностями, не являются статическими образованиями, но могут изменяться в процессе индивидуального развития организма, что, в свою очередь, сопровождается изменением характера экспрессии больших блоков генов в дифференцирующихся клетках.

Структурно-функциональный анализ инсуляторов в вышеупомянутой системе с геном *white* показал, что последовательности участков инсуляторов, гиперчувствительных к ДНКазе I, необходимы для их функционирования, тогда как центральная А/Т-богатая область, устойчивая к действию нуклеазы, для этого не требуется. Удаление части последовательностей из гиперчувствительных сайтов сопровождается понижением активности инсуляторов, тогда как простое увеличение числа копий последовательностей восстанавливает их активность. На этом основании делается вывод, что активность пограничных последовательностей может обеспечиваться определенным критическим количеством молекул связавшихся с ними белков, неспецифичных в отношении типа инсулятора, которые могут действовать на энхансеры либо непосредственно, либо через изменение структуры хроматина.

Природа белкового компонента, взаимодействующего с пограничными последовательностями, была определена в случае типичного инсулятора: *scs'*-последовательности гена теплового шока дрозофилы. Из ядер культивируемых клеток был выделен белок BEAF-32 (boundary element associated factor) с молекулярной массой 32 кДа, который взаимодействует с палиндромной последовательностью, фланкирующей два гиперчувствительных к ДНКазе участка в *scs'*. С помощью иммунохимических методов локализовали множественные высокоспецифические места связывания BEAF-32 на политенных хромосомах с *одной из сторон* пуффов, образующихся на определенных стадиях развития личинок мух, что подчеркивает двухкомпонентный состав последовательностей нуклеотидов инсуляторов. Хотя белки, связывающиеся с противоположной частью этой пограничной последовательности, еще не идентифицированы, предполагают, что белки двух частей инсулятора взаимодействуют друг с другом. Полагают также, что такого рода связывание белков, ассоциированных с различными инсуляторами, может быть одним из необходимых условий формирования большого числа разнообразных функционально активных доменов хромосом, что, в свою

очередь, может обеспечивать специфическую экспрессию заключенных в них генов на разных стадиях онтогенетического развития организма.

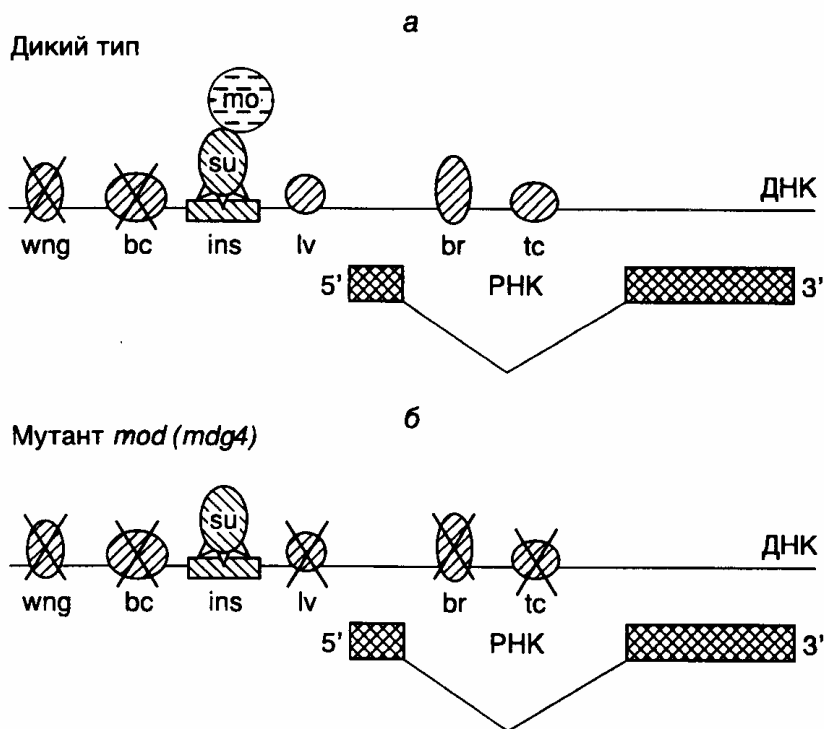


Рис. 1.34. Структура и функционирование инсультатора ретротранспозона *gypsy* дрозофилы

а – последовательность инсультатора (*ins*), представленная 12 копиями сайта связывания белка *su(Hw)* (*su*), взаимодействующего с белком *mod(mdg4)* (*mo*). Комплекс этих белков, связанных с инсультатором, односторонне подавляет активность энхансеров (перечеркнутые заштрихованные эллипсы) гена *yellow* в соответствующих тканях; **б** – инактивация белка *mod(mdg4)* под действием мутации разрушает белковый комплекс *su*–*mo*, и подавление транскрипции под действием инсультатора становится неспецифическим (двусторонним). Обозначения тканей дрозофилы: *wng* – крылья, *bc* – катикла тела, *lv* – ткани личинки, *tc* – тарзальные коготки, *br* – щетинки

Вышеупомянутый инсультатор ретротранспозона *gypsy* содержит 12 сайтов связывания белка *su(Hw)*. Коровая последовательность этих сайтов гомологична октамерной последовательности, обнаруженной в различных энхансерах и промоторах позвоночных. Эта последовательность фланкирована АТ-богатыми участками, которые способствуют изгибанию молекулы ДНК, требуемому для правильных ДНК-белковых взаимодействий в этом участке

генома. Полипептидная цепь белка *su(Hw)* содержит 12 доменов типа "цинковые пальцы", необходимых для его связывания с ДНК и функционирования инсулятора. Кроме того, белок *su(Hw)* обладает двумя кислыми доменами, локализованными вблизи его С-конца, которые обеспечивают его взаимодействие с энхансерами и нужны для подавления эффекта положения. Абсолютно необходимым для функционирования инсуляторов является и α -спиральный участок полипептидной цепи этого белка, гомологичный второму спиральному участку (helix-coiled coil) основных факторов транскрипции группы HLH-bzip, содержащему домен типа "лейциновая застежка". Поскольку такой домен обычно требуется для осуществления белок-белковых взаимодействий, полагают, что для функционирования этого инсулятора кроме белка *su(Hw)* необходимы и другие белки.

С помощью классического генетического анализа был идентифицирован второй компонент инсулятора *su(Hw)*. Мутации в гене *modifier of mdg4* (*mod(mdg4)*) ингибировали полярное действие инсулятора на энхансеры (рис. 1.34,а) и усиливали мозаичный эффект положения для генов, транслоцированных в гетерохроматин. У таких мутантов действие инсулятора на подавление активности энхансеров было двунаправленным и не зависело от расстояния между энхансерами и промотором (см. рис. 1.34,б). Ген *mod(mdg4)* кодирует, по крайней мере, три белка, которые возникают в результате альтернативного сплайсинга и содержат ВТВ-домен, характерный для многих факторов транскрипции, например уже упоминаемого GAGA-фактора, осуществляющего свое действие через изменение структуры хроматина. Полагают, что двунаправленное подавление функций энхансеров инсулятором *gypsy* в отсутствие функционально активного белка *mod(mdg4)* являются следствием гетерохроматизации последовательностей, окружающих инсулятор. В этой связи подавление активности энхансеров у мутантов *mod(mdg4)* может происходить из-за изменений в нуклеосомной структуре данных участков ДНК. Поскольку для развития эффекта требуется наличие нативных доменов типа "лейциновая застежка" и кислых доменов белка *su(Hw)*, предполагают, что процесс гетерохроматизации ДНК в отсутствие функционального белка *mod(mdg4)* является следствием взаимодействия белка *su(Hw)* с другими белками. Ингибирующее действие инсулятора ретротранспозона *gypsy* может

распространяться на энхансеры, находящиеся по отношению к нему в *транс*-положении, т.е. на гомологичной хромосоме.

Дальнейшим указанием на то, что действие пограничных последовательностей на соответствующие гены осуществляется через изменение структуры хроматина, было получено при исследовании их влияния на дозовую компенсацию у дрозофилы. Для экспрессии генов, расположенных на X-хромосоме самцов и самок этого организма, характерен одинаковый уровень, хотя содержание (доза) таких генов у самок в два раза выше, чем у самцов. Это равновесие достигается за счет механизма, повышающего вдвое скорость транскрипции сцепленных с X-хромосомой генов у самцов и получившего название дозовой компенсации. На X-хромосоме самцов происходит сборка мультимерных белковых комплексов (так называемых *летальных комплексов*), которые изменяют нуклеосомную структуру хромосомы за счет ацетилирования остатков Lys-16 гистона H4. Способность к дозовой компенсации полностью отсутствует у локусов X-хромосомы, перенесенных на аутосомы, т.е. последние подавляют этот эффект. Однако если трансген фланкирован пограничными последовательностями ретротранспозона *gypsy*, то у 90% трансгенов, интегрированных в аутосомы, происходит правильная дозовая компенсация. Механизм этого явления пока непонятен. Предполагают, что изменение структуры хроматина в трансгене под действием инсуляторов делает возможным сборку на нем летального комплекса или может препятствовать доступу деацетилаз к гистонам нуклеосом.

Результаты исследований белков BEAF-32, su(Hw) и mod(mdg4) указывают на то, что инсуляторы и пограничные последовательности формируют мультибелковые комплексы, регулирующие экспрессию генов через изменение структуры соседнего с ними хроматина. Такие изменения конформации хроматина могут оказывать влияние на взаимодействия между энхансерами и промоторами генов, не изменяя функциональной активности энхансеров как таковых. При этом изменения структуры хроматина, вызываемые инсуляторами, не препятствуют элонгации РНК РНК-полимеразами.

В настоящее время идентифицированы и другие хромосомные белки, обеспечивающие инактивацию генов через конденсацию хроматина, которые,

кроме того, необходимы для поддержания хроматина и целых хромосом в конденсированном неактивном состоянии. К таким белкам, в частности, относится белок гетерохроматина 1 (HP-1) дрозофилы, ассоциированный с β -гетерохроматином. Мутации в гене, кодирующем белок HP-1, являются супрессорными, подавляющими мозаичный эффект положения (PEV). В то же время дупликация этого локуса усиливает PEV, что указывает на зависимость упаковки гетерохроматина, инактивирующей гены, от внутриклеточной концентрации белка HP-1. Для такого белка характерна высокая эволюционная консервативность, и его гомологи обнаружены у мышей, человека и растений. Использование переходов конденсации–деконденсации хроматина для регуляции экспрессии как отдельных генов, так и их громадных массивов, по-видимому, является прерогативой эукариот. Еще одним ярким примером такого рода служит инактивация одной из X-хромосом самок млекопитающих в раннем эмбриогенезе.

Инактивация X-хромосом. Поскольку соматические клетки самок млекопитающих содержат две половые X-хромосомы, а самцов – только одну, у самок возникает необходимость в компенсации двойной дозы генов, сцепленных с этими хромосомами. Такая проблема решается путем инактивации одной из X-хромосом, являющейся следствием перевода ее хроматина в высококонденсированное состояние с подавленной транскрипцией. Хотя в большинстве случаев в разных клетках одного и того же животного случайным образом инактивируется или материнская, или отцовская X-хромосома, имеются специальные случаи, когда во всех клетках только отцовская хромосома становится транскрипционно неактивной. Последний случай, получивший название "импринтный тип инактивации X-хромосомы", характерен для сумчатых животных и обнаружен в некоторых клетках эмбрионов мышей. (Более подробно о генетическом импринтинге см. раздел 3.2.5.)

Лучше всего молекулярные механизмы инактивации X-хромосомы изучены у человека, мышей и сумчатых животных. В ранних эмбрионах мышей и человека обе X-хромосомы активны. У мышей инактивация X-хромосомы впервые обнаруживается в ранних эмбрионах на стадии бластоцисты, когда для всех внеэмбриональных клеток характерен импринтный тип инактивации только отцовских X-хромосом. Позднее, через 6 дней эмбрионального развития

во всех клетках зародыша материнские и отцовские X-хромосомы инактивируются случайным образом. Неактивное состояние хромосом высокостабильно, и их реактивация (повторная активация) имеет место лишь у самок в клетках зародышевой линии. В клетках зародышевой линии самцов единственная X-хромосома инактивируется в раннем сперматогенезе.

Свойства инактивированных X-хромосом. Все инактивированные X-хромосомы обладают некоторыми общими свойствами. Их ДНК реплицируется в поздней S-фазе клеточного цикла, а в интерфазе они присутствуют в высококонденсированном состоянии в виде так называемых *телец полового хроматина* (sex chromatin bodies). Все это происходит на фоне низкого уровня ацетилирования гистонов и появления нового гистона масо H2A1, что указывает на возможные изменения структуры нуклеосом. В эмбриональных соматических клетках (но не в клетках зародышевой линии) плацентарных млекопитающих метилированы CpG-островки генов инактивированной X-хромосомы, в отличие от CpG-островков остальной части генома. Неметилированными остаются и CpG-островки инактивированных X-хромосом сумчатых животных. Большая часть вышеперечисленных свойств инактивированных X-хромосом присуща и обычному гетерохроматину. Поэтому процесс инактивации рассматривают как гетерохроматизацию X-хромосом.

У редко встречающихся индивидуумов с ненормальным числом X-хромосом – X0, XX, XXY и т.п. независимо от их числа – по-прежнему инактивированной остается только одна половая хромосома. В отличие от этого, в клетках триплоидных организмов, содержащих по три набора хромосом, инактивируется или одна, или две X-хромосомы, а у тетраплоидов – две. Такого рода данные указывают на существование механизма, поддерживающего в активном состоянии одну X-хромосому на каждые два набора аутом. Таким образом, перед инактивацией в эмбриональных клетках происходит определение числа наборов аутом, выбор X-хромосом, остающихся активными, и инициация инактивации остальных X-хромосом.

Инициация инактивации X-хромосом. Процесс инактивации X-хромосом начинается на определенном участке, получившем название *центра инактивации X-хромосомы* (X-inactivation center). В этом участке был картирован ген *Xist* (X inactive specific transcript), который впоследствии был клонирован. Сегменты X-хромосомы, у которых центр инактивации удален в

результате делеций или транслокаций, не подвергаются инактивации. С помощью генного нокаута *Xist* было установлено, что для инактивации X-хромосом требуется его транскрипция. Введение гена *Xist* с помощью трансгеноза в аутосомы было достаточным для инициации инактивации последних.

Ген *Xist* как основной регулятор инактивации X-хромосом. С помощью генного нокаута удалось показать, что за определение внутриклеточного числа X-хромосом (counting) и инициацию их инактивации отвечают разные участки гена *Xist*. 5'-Концевая часть гена необходима для инициации инактивации, а его 3'-концевая зона участвует в счете хромосом, в результате которого определяется, какая хромосома остается активной. X-Хромосомы с поврежденной 3'-концевой частью гена *Xist* всегда инактивируются. В этой серии экспериментов было также установлено, что ген *Xist* требуется как для случайной инактивации хромосом, так и для инактивации импринтного типа, однако его участие, возможно, не требуется для инактивации X-хромосом в мужских зародышевых клетках. В связи с тем, что на активных X-хромосомах ген *Xist* сильно метилирован, полагают, что его метилированное состояние сопровождается репрессией гена, препятствующей инактивации X-хромосом.

Удивительным свойством гена *Xist* является то, что он не кодирует белки. В результате экспрессии гена образуется большая РНК, длина которой составляет 15–17 т.о. После окончания синтеза РНК остается в ядре, а в клетках, где X-хромосома уже инактивирована, эта РНК создает "оболочку" вокруг неактивной хромосомы. В недифференцированных эмбриональных стволовых клетках, где обе X-хромосомы активны, зоны экспрессии гена *Xist* выявляются с помощью гибридизации *in situ* в виде отдельной точки на каждой хромосоме на фоне общего низкого содержания *Xist*-РНК. Во время дифференцировки клеток зона содержания *Xist*-РНК расширяется вдоль будущей неактивной хромосомы на фоне увеличения ее количества. Одновременно резко уменьшается уровень экспрессии гена *Xist* на хромосоме, остающейся активной.

Повышение внутриклеточного уровня *Xist*-РНК происходит не за счет стимуляции транскрипции гена, но вследствие стабилизации самих молекул РНК. При этом стабильная и нестабильная РНК транскрибируются с разных

промоторов одного и того же гена. И стабилизация, и хромосомная локализация Xist-PHK являются необходимыми, но не достаточными условиями инактивации X-хромосомы. В ранних эмбрионах мышей на стадии предимплантации отцовские X-хромосомы покрыты оболочкой Xist-PHK, однако остаются в активном состоянии. Это указывает на участие дополнительных факторов в эмбриональной инактивации X-хромосом, которые отсутствуют в клетках ранних эмбрионов.

Установление и поддержание неактивного состояния X-хромосом.

Механизм распространения зоны неактивного хроматина вдоль X-хромосом от центра инактивации пока остается непонятным. В аутосомах, содержащих транслоцированный центр инактивации или интегрированный трансген *Xist*, инактивация распространяется в обе стороны на многие миллионы пар оснований. Однако в этом случае гетерохроматизация хроматина происходит менее эффективно и по-разному на индивидуальных хромосомах. При этом некоторые гены аутосом остаются в активном состоянии. Полагают, что для X-хромосом характерно наличие специфических последовательностей, необходимых для их эффективной гетерохроматизации при инактивации, которые отсутствуют в аутосомах.

Некоторые гены инаktivированных X-хромосом продолжают функционировать. Часть из них имеет гомологов на Y-хромосоме и поэтому не требует дозовой компенсации. В ряде случаев их активированное состояние является следствием реактивации предварительно инаktivированного хроматина. Однако для большинства таких генов механизм поддержания в активном состоянии на фоне молчащего генетического окружения остается неизвестным.

Ген *Xist* транскрибируется в клетках всех женских особей в присутствии метилированных CpG-островков генов инаktivированных X-хромосом. В культивируемых клетках деметилирующие агенты вызывают частичную реактивацию молчащих генов на этих хромосомах. Не исключено, что кроме этих механизмов инаktivированное состояние X-хромосом поддерживает и поздняя репликация их ДНК. У сумчатых животных имеет место поздняя репликация, однако не наблюдается дифференциального метилирования CpG-островков, и у них инаktivированное состояние X-хромосом менее стабильно, чем у плацентарных млекопитающих. В последнем случае даже прекращение

транскрипции гена *Xist* не сопровождается реактивацией молчащих X-хромосом, однако на фоне слабого метилирования такая реактивация иногда происходит. Таким образом, поддержание инактивированного состояния X-хромосомы является сложным многоуровневым процессом.

Импринтная инаktivация только отцовских X-хромосом рассматривается как эволюционно более примитивная форма дозовой компенсации у животных. Действительно, в этом случае повышается вероятность вредного воздействия на организм мутантных материнских генов, локализованных на X-хромосоме. Такая вероятность должна быть меньше при случайной инаktivации в разных клетках отцовских и материнских хромосом. Наличие молекулярных механизмов, регулирующих экспрессию генов через дозовую компенсацию, позволило высказать предположение, что эволюционный обмен генами между аутосомами и половыми хромосомами является вредным для биологического вида, так как может приводить к регуляторному дисбалансу экспрессии транслоцированных генов. Был сформулирован так называемый закон Оно, в соответствии с которым гены, сцепленные с X-хромосомой у одного вида млекопитающих, должны оставаться сцепленными с X-хромосомой и у всех остальных.

Аналогично тому, что было показано на примере X-хромосомы сумчатых млекопитающих, при делении во время митоза дочерние клетки могут наследовать от родительских не только прямую генетическую информацию в виде новой копии всех генов, но и определенный уровень их активности (в данном примере – активное или неактивное состояния). Такой тип наследования генетической информации получил название *эпигенетического наследования*. Современная эпигенетика изучает наследуемые особенности (паттерны) экспрессии генов, вызываемые потенциально обратимыми изменениями структуры хроматина и(или) метилирования ДНК, не сопровождаемые изменениями ее первичной структуры.

3.2.5. Импринтинг

Другим характерным примером регуляции экспрессии генов, приводящей к эпигенетическому наследованию признаков, является уже упомянутый выше импринтинг, при котором специфический характер дифференциальной активности генов определяется полом организма, от которого эти гены

унаследованы. В частности, у некоторых насекомых, например грибных комариков (Sciaridae), весь набор отцовских хромосом элиминируется во время сперматогенеза. У этих организмов отцовские хромосомы маркируются в цитоплазме клеток зародышевой линии, удаляются при созревании гамет и не передают свои гены следующему поколению. У млекопитающих и высших цветковых растений отцовские и материнские гены оказывают разный эффект на развитие эмбриона, но в одинаковой степени представлены в гаметах, образующихся в результате мейоза. В этом случае вклады отцовского и материнского геномов в развитие организмов не эквивалентны и происходит видимое искажение менделевских правил наследования некоторых признаков. Иными словами, характер экспрессии отцовского и материнского аллелей одного и того же гена в организме-потомке может быть различным, так как зависит от их происхождения. Молекулярные механизмы этого явления в настоящее время до конца непонятны. Предполагается, что в развитие феномена импринтинга вносят вклад зависимые от пола модификации ДНК определенных генов, в частности метилирование их регуляторных участков.

3.2.6. Метилирование ДНК в регуляции транскрипции

Единственной известной генетически запрограммированной ковалентной модификацией ДНК у высших эукариот является метилирование остатков цитозина в положении 5 с образованием 5-метилцитозина (5-mC). Эта реакция катализируется ферментом (цитозин-5)-ДНК-метилтрансферазой (Мтазой), который обнаружен у прокариот и эукариот. Каталитический механизм действия этих ферментов был подробно изучен на примере бактериальной метилазы *M.HhaI*, которая модифицирует соответствующий сайт рестрикции. С помощью рентгеноструктурного анализа было показано, что после взаимодействия Мтазы со специфическим участком ДНК остаток модифицируемого цитозина выпячивается из двойной спирали и образует ковалентную связь с полипептидной цепью фермента в положении С6. В этом промежуточном комплексе активированный атом углерода С5 акцептирует метильную группу S-аденозилметионина, выступающего в качестве кофактора. Полипептидные цепи Мтаз эукариот содержат на своих N-концах большой домен, который обеспечивает ядерную локализацию ферментов, их доставку к репликативным вилкам, ответ ферментов на регуляторные воздействия.

Большинство прокариотических Мтаз способны метилировать ДНК *de novo*, распознавая неметилированные палиндромные гексануклеотидные последовательности. Они также метилируют последовательности, в которых одна цепь ДНК уже содержит метильные группы. В отличие от этого эукариотические Мтазы относятся к "поддерживающим" ферментам, которые узнают и метилируют только наполовину метилированные последовательности, формирующиеся во время репликации ДНК, когда вновь синтезированная цепь неметилирована. У млекопитающих остатки С метилируются преимущественно в составе динуклеотидов CpG, однако в последнее время описаны случаи метилирования и последовательностей CpNpG. В геноме позвоночных животных метилировано ~70% динуклеотидов CpG и ~6–7% всех остатков цитозина.

"Поддерживающие" Мтазы животных обладают небольшой способностью осуществлять метилирование ДНК и *de novo* в полностью неметилированных участках, а также искусственных субстратов (олигонуклеотидов), содержащих ошибочно спаренные основания. Остается непонятным, является ли указанное свойство Мтаз достаточным для осуществления метилирования *de novo* обширных участков генома в эмбриогенезе или же этот процесс происходит с участием других ферментов. Известно, что гомозиготные делеции в гене Мтазы у мышей вызывают гибель зародышей в раннем эмбриогенезе, что указывает на важную роль метилирования ДНК в онтогенезе млекопитающих. Однако даже у таких мутантных эмбрионов небольшая часть последовательностей ДНК метилирована.

Метилирование остатков цитозина оказывает влияние на структурные характеристики ДНК. Это проявляется в облегчении перехода метилированных участков ДНК из В-формы в Z-форму, увеличении шага спирали ДНК и изменении кинетики образования крестообразных структур. Метильная группа 5-мС выступает на поверхности большой бороздки ДНК, находящейся в В-форме, и увеличивает ее гидрофобность, что в ряде случаев является решающим фактором при взаимодействии белков с соответствующими участками ДНК.

Регуляция экспрессии генов с помощью метилирования ДНК.

Метилирование остатков С может оказывать влияние на транскрипцию как непосредственно через изменение эффективности связывания позитивных и

негативных факторов транскрипции со своими регуляторными участками на ДНК, так и опосредованно через формирование неактивных в транскрипционном отношении участков хроматина. Поскольку 5-мС структурно подобен тимину, метилирование остатков С может сопровождаться возникновением новых консенсусных последовательностей для некоторых факторов транскрипции. В частности, метилирование превращает низкоаффинный сайт связывания фактора транскрипции AP-1 (CGAGTCA) в высокоаффинный сайт (mCGAGTCA), который соответствует консенсусному сайту для этого фактора (TGAGTCA). Также неоднократно описано ингибирование взаимодействия некоторых белков с ДНК в результате метилирования CpG-последовательностей в сайтах их связывания (табл. I.15).

Таблица I.14

Факторы транскрипции позвоночных, на активность которых оказывает влияние метилирование остатков цитозина в узнаваемых ими регуляторных последовательностях нуклеотидов

Фактор транскрипции	Узнаваемая последовательность
AP-2	CCCGC <u>CG</u> GCG
AP-2	CTCC <u>CG</u> GGG(C/T)TG
Ah-рецептор	TTG <u>CG</u> TG
CREB/ATF	TGAC <u>CG</u> TCA
CREB/ATF	CTG <u>CG</u> TCA
E2F	TTTTCG <u>CG</u> C
EBP-80	ATCTG <u>CG</u> CATATGCC
Ets	<u>CG</u> GAAG
MLTF	GGCCAC <u>CG</u> TGACC
MTF-1	TG <u>CG</u> CC <u>CG</u>
c-Myc, c-Min	CAC <u>CG</u> TG
GABP	TTCC <u>CG</u> GGC
NF-κB	GGGACTTT <u>CCG</u>
HiNF-P	<u>CG</u> GTTTTCAATCTGGT <u>CCG</u>
MSPF	ATG <u>CG</u> NNNN <u>CG</u> CCT

В табл. I.15 представлены только консенсусные последовательности, узнаваемые соответствующими факторами транскрипции с подчеркнутыми динуклеотидами CpG, метилирование которых оказывает влияние на взаимодействие факторов с ДНК. Интересно, что фактор Sp1, необходимый для инициации транскрипции на промоторах многих генов домашнего хозяйства, может связываться с метилированными и неметилированными консенсусными последовательностями и, кроме того, предотвращать метилирование соседних последовательностей нуклеотидов. Поскольку Sp1 обнаружен у высших эукариот, ДНК которых содержит 5-мС, предполагают, что данный фактор может играть специфическую роль в регуляции метилирования ДНК у этих организмов.

Известные промоторы, за небольшим исключением (например промотора гена главного комплекса гистосовместимости H-2K), неактивны в метилированном состоянии. Единственная метильная группа, введенная в промотор гена тимидинкиназы вируса простого герпеса или в С-промоторы вируса Эпштейна–Барр, может оказывать большое негативное влияние на транскрипцию. Уровень подавления активности других промоторов, в частности промоторов α - и γ -глобиновых генов человека или промотора мышинного гена MyoD1, находится в прямой зависимости от числа введенных в них метильных групп, но не от их положения в промоторах. Ингибирование транскрипции, вызванное частичным метилированием промоторов, может преодолеваться с помощью энхансеров, однако полностью метилированные промоторы не реактивируются энхансерами и сохраняют свое репрессированное состояние, несмотря на присутствие последних. На основании такого рода данных высказывается предположение, что метилирование ДНК регулирует транскрипцию по принципу "все или ничего" и не обеспечивает тонкой регуляции экспрессии генов.

Влияние метилирования ДНК на структуру хроматина. Характер метилирования ДНК может оказывать решающее влияние на структуру хроматина и фазирование (phasing) нуклеосом. Термином "фазирование" обозначают неслучайное расположение нуклеосом относительно конкретной последовательности нуклеотидов ДНК в определенных участках генома. Большинство нуклеосом генома животных располагаются на ДНК случайно. В редких случаях фазирования последовательности нуклеотидов упаковываются

в нуклеосомы одним и тем же способом во всех клетках организма, что может оказывать влияние на эффективность транскрипции и репликации ДНК в соответствующих локусах. В настоящее время у млекопитающих описано несколько белков, которые взаимодействуют преимущественно с метилированной ДНК. К ним относятся, например вирусный белок MDBP1 (Major DNA-Binding Protein) и белки человека MeCP1/2 (Methyl-CpG-Binding Protein), а также комплекс белков MDBP2 и гистона H1. Если белок MeCP2 для своего взаимодействия требует присутствия в соответствующем участке ДНК только одного метилированного CpG, то белок MeCP1 связывается с ДНК лишь при наличии нескольких метилированных динуклеотидов. С использованием специфических антител было установлено, что белок MeCP2 локализован преимущественно в G-полосах гетерохроматина метафазных хромосом. Гомозиготные делеции в гене этого белка летальны для мышей, что указывает на важную роль MeCP2 в эмбриогенезе животных. Комплекс MDBP2–гистон H1 может связываться с участком ДНК, содержащим единственный метилированный CpG, и ингибировать транскрипцию. При этом белково-нуклеиновое взаимодействие усиливается после фосфорилирования белков.

Ярким примером подавления экспрессии генов путем формирования хроматина, структурированного специфическим образом, является уже обсуждавшаяся выше инактивация одной из X-хромосом самок млекопитающих. Большинство генов неактивной X-хромосомы прекращает свою экспрессию в раннем эмбриогенезе независимо от наличия в ней CpG-последовательностей. При этом метилирование ДНК и ацетилирование гистона H4 усиливают и стабилизируют репрессированное состояние неактивной X-хромосомы.

Активация генов, входящих в состав неактивного хроматина, требует изменения структуры хроматина и деметилирования ДНК. Обнаружены белковые комплексы, способные реактивировать неактивный хроматин. Деметилирование может происходить пассивно через подавление функционирования поддерживающих Mtaз. В этом случае дочерние цепи ДНК, образующиеся в процессе репликации, остаются неметилированными, что приводит к полному деметилированию ДНК после двух раундов репликации. Активное деметилирование ДНК происходит с участием специфических ферментов – деметилаз.

Гены, входящие в состав неактивного хроматина, в большинстве своем недоступны действию факторов транскрипции. Формирование неактивного хроматина через метилирование специфических последовательностей приводит к уменьшению линейных размеров генома и числа регуляторных последовательностей, доступных факторам транскрипции. Следствием этого является снижение неспецифических взаимодействий регуляторных белков с соответствующими последовательностями генов, что, в свою очередь, уменьшает уровень информационного шума в генетической системе, который может появляться в результате неконтролируемой транскрипции. В этой связи высказывается предположение о важной роли не только генетических, но и видоспецифических эпигенетических изменений в эволюции высших эукариот. Действительно, наследуемый характер паттернов метилирования геномной ДНК способствует стабилизации и распространению в популяциях специфических структур хроматина, а изменение паттернов через эпигенетические мутации может быть одним из путей повышения пластичности генетической информации и ее разнообразия, что может служить материалом для естественного отбора.

СрG-островки. Для генома эукариот характерно наличие *изохор* (isochores) – протяженных последовательностей, обогащенных GC- или AT-парами. GC-богатые изохоры локализуются в рано реплицирующихся участках хромосом и ассоциированы с R-сегментами (полосами), выявляемыми при их дифференциальной окраске. AT-богатые изохоры реплицируются в поздней S-фазе клеточного цикла и часто обнаруживаются в G-сегментах хромосом. СрG-Островки (CpG islands) представляют собой GC-богатые последовательности (60–70% GC) длиной не менее 200 п.о. с большим числом динуклеотидов CpG. Они локализуются преимущественно в R-сегментах и поддерживаются в неметилованном состоянии в клетках зародышевой линии и нормальных соматических тканей. Исключение составляют островки в генах, подвергнутых импринтингу, в генах инактивированных X-хромосом, а также в повторяющихся элементах генома LINE и SINE (например Alu), где они сильно метилированы. Полагают, что именно неметилованное состояние СрG-островков стабилизирует их структуру в эволюции генома из-за более низкой скорости мутирования по сравнению с 5-mC, который в результате дезаминирования часто превращается в тимин со всеми вытекающими отсюда мутагенными

последствиями. В гаплоидном геноме человека обнаруживают ~ 45000 CpG-островков, и считается, что большинство из них ассоциировано с генами. Промоторы многих генов домашнего хозяйства содержат неметилированные CpG-островки.

В настоящее время непонятен механизм поддержания CpG-островков в неметилированном состоянии. Поскольку Мтаза обладает способностью метилировать их *in vitro*, структурные особенности ДНК сами по себе не способны это обеспечивать. Считается, что в поддержании неметилированного состояния островков участвуют специфические белки хроматина. Хроматин CpG-островков не содержит гистона H1, а гистоны H2A и H2B в этих участках хроматина недоацетилированы. Кроме того, CpG-островки часто содержат сайты связывания фактора транскрипции Sp1, который, взаимодействуя как с метилированными, так и неметилированными сайтами, предотвращает их исходное или дальнейшее метилирование. Полагают, что именно этот фактор может играть ключевую роль в постоянном поддержании генов домашнего хозяйства в неметилированном активном состоянии.

Для соматических клеток, растущих в первичной культуре, характерно быстрое изменение характера метилирования ДНК. Это изменение часто связывают с ограниченным пролиферативным потенциалом культивируемых клеток и их старением как *in vivo*, так и *in vitro*. Многие гены культивируемых клеток гиперметилированы, а отдельные CpG-островки аутосом, поддерживаемые *in vivo* в неметилированном состоянии, переходят в метилированное состояние. Полагают, что гиперметилирование ДНК и инактивация многих генов культивируемых клеток являются следствием отбора, направленного на выключение экспрессии генов, которые не требуются безусловно для выживания клеток в культуре. Кроме того, в культуре происходит отбор клеток на способность к быстрому делению, поскольку быстро делящиеся клетки вытесняют медленно пролиферирующие. Поэтому гены, вызывающие задержку клеточного цикла и терминальную дифференцировку клеток, например p16 или MyoD1, быстро инактивируются у культивируемых клеток под действием мутаций или гиперметилированием.

Роль метилирования ДНК в дифференцировке клеток. Для разных органов и тканей организма характерны специфические наборы (паттерны) метилированных последовательностей ДНК. Такой мозаицизм в метилировании

последовательностей генома не является следствием одних только ошибок функционирования системы метилирования, поскольку паттерны метилирования ДНК в гомологичных тканях разных индивидуумов обладают большим сходством. В этой связи можно говорить о тканеспецифическом характере метилирования ДНК, которое обеспечивается координированным функционированием системы метилаз и деметилаз в онтогенезе организма.

В экспериментах на культурах клеток было показано, что изменение паттернов метилирования под действием веществ, вызывающих деметилирование ДНК, например 5-азацитидина, может изменять дифференцированное состояние клеточных линий. В частности, инкубация эмбриональных фибробластов мышей с 5-азацитидином сопровождается их превращением в миобласты, адипоциты и хондроциты. Точно так же дифференцировка клеток эритролейкемии Френд может быть запущена с помощью деметилирующих агентов. Полагают, что в этих случаях в процессе деметилирования ДНК происходит активация генов-регуляторов, детерминирующих дифференцированное состояние клеток, которые, в свою очередь, индуцируют экспрессию всех остальных генов, необходимых для поддержания дифференцированного состояния. Экспрессия конкретных доминантных генов, ответственных за дифференцировку отдельных групп клеток в определенном направлении, должна жестко контролироваться и быть блокированной в клетках других типов. Метилирование ДНК может обеспечивать такое установление неактивного состояния генов дифференцировки и его поддержание в ряду клеточных поколений, т.е. контролировать стабильность дифференцированного состояния соматических клеток организма.

Отцовский импринтинг и аллельное исключение. При отцовском импринтинге, как об этом уже упоминалось выше, экспрессия некоторых генов в дочерних соматических клетках зависит от того, локализованы ли они на отцовских или материнских хромосомах. При этом часто наблюдают различный характер метилирования отцовских и материнских хромосом. У мышей известны ~15 областей хромосом, подвергнутых импринтингу. У многих генов с таким эпигенетическим наследованием обнаружены прямые повторы, которые, как полагают, могут быть задействованы в установлении и поддержании импринтинга. Локализация нескольких генов, в регуляции экспрессии которых

задействован импринтинг, в одном и том же участке хромосомы 16 указывает на то, что в установление импринтинга могут быть вовлечены целые домены хромосом.

При *аллельном исключении* происходит инактивация одного из аллелей, приводящая к функциональной гемизиготности соответствующего гена независимо от его происхождения со стороны родителей. Под гемизиготностью диплоидных организмов или отдельных соматических клеток понимают представленность одного или нескольких генов только одним аллелем, что может быть следствием анеуплоидии (потери одной или нескольких хромосом), делеции или функциональной инактивации второго аллеля. Гемизиготное состояние характерно, в частности для генов X-хромосом самок млекопитающих отдельных соматических клеток, где одна из X-хромосом инактивирована случайным образом. Феномен аллельного исключения описан для генов иммуноглобулинов животных, лишь один из аллелей которых претерпевает продуктивную перестройку под действием V-D-J-рекомбиназы в отдельных лимфоцитах развивающегося организма. Метилирование ДНК оказывает влияние на объединение V-D-J-сегментов генов и, возможно, играет ключевую роль в их аллельном исключении.

Обычно аллельное исключение тонко регулируется в онтогенезе. Однако теоретически аналогичный механизм может реализовываться в результате постепенно накапливаемых случайных ошибок метилирования на протяжении жизни организма. Накопление ошибок метилирования сверх определенного уровня должно сопровождаться выключением соответствующего аллеля на фоне нормального функционирования другого аллеля в конкретной соматической клетке. Происхождение часто обнаруживаемых частично метилированных CpG-сайтов в ДНК различных тканей объясняют такими ошибками метилирования, получившими название эпимутаций. На фоне эпимутаций может легко произойти инактивация единственного оставшегося аллеля дикого типа под действием обычных мутаций или в результате гиперметилирования, что приведет к полному подавлению экспрессии соответствующего гена.

Метилирование как механизм инактивации чужеродной и повторяющейся ДНК у эукариот. Т. Бестором (1990 г.) было высказано предположение о том, что метилирование ДНК у эукариот может выполнять

защитную функцию, предохраняя организм от чужеродной ДНК и избытка эндогенных повторяющихся последовательностей. Поскольку у эукариот не описано ферментов рестрикции, специфичных в отношении GrC-последовательностей, предполагается, что такие защитные механизмы осуществляют инактивацию чужеродной ДНК, а не ее деградацию, как это имеет место у бактерий. Экспериментально установлено, что ДНК из внешних источников, например вирусная ДНК или ДНК, введенная в клетки с помощью трансфекции, а также эндогенная ДНК ретротранспозонов и повторяющихся последовательностей (в частности LINE и SINE) могут быть объектом метилирования *de novo*, приводящего к их генетической инактивации. Например, интеграция ДНК аденовирусов в эукариотический геном сопровождается ее постепенным метилированием. То же характерно и для латентного состояния некоторых других вирусов, которое поддерживается через метилирование остатков цитозина в ДНК их геномов Мтазами клетки-хозяина. Другим примером функционирования этого механизма является гиперметилирование трансгенов, интегрированных в геном организма-реципиента в большом числе копий.

Механизмы, которые позволяют высшим эукариотическим организмам избирательно распознавать и инактивировать чужеродную и повторяющуюся ДНК, не совсем понятны. Полагают, что Мтазы преимущественно распознают молекулы ДНК, формирующие необычные пространственные структуры, например петли или ошибочно спаренные нуклеотиды. В случае повторяющихся последовательностей мишенями Мтаз при метилировании *de novo* могут быть интермедиаты, образующиеся во время их рекомбинации друг с другом. У гриба *Ascobolus immersus* повторяющиеся последовательности эффективно распознаются и инактивируются в результате метилирования под действием механизма, получившего название премейотически индуцированного метилирования (methylation induced premeiotically – MIP). Аналогичный механизм у гриба *Neurospora crassa* назван мутагенезом, индуцированным повторяющимися последовательностями (repeat induced point mutations – RIP). В этом случае остатки цитозина в повторах либо метилируются ферментными системами гриба, либо превращаются в мутантные остатки тимина. Вообще, остатки С, которые являются мишенями для Мтаз, мутируют в геноме эукариот по механизму дезаминирования и

превращения в остатки тимина гораздо чаще, чем другие остатки цитозина. Это позволяет рассматривать метилирование как постоянно функционирующий механизм эукариот, инактивирующий чужеродные и повторяющиеся последовательности ДНК с помощью таких мутаций.

Все рассмотренные выше регуляторные механизмы, контролирующие экспрессию генов, функционируют на одном из самых важных этапов реализации генетической информации – на уровне транскрипции. Однако и другие этапы передачи генетической информации являются важными объектами многочисленных регуляторных воздействий.

3.3. Посттранскрипционная регуляция экспрессии генов

Регуляция экспрессии генов на уровне уже синтезированной РНК играет исключительно важную роль в онтогенезе многоклеточных и особенно высших организмов. Среди различных способов посттранскрипционной регуляции наибольшее значение имеют механизмы внутриклеточной локализации и депонирования РНК с последующим специфическим вовлечением ее в трансляцию, а также разные формы процессинга предшественников мРНК, избирательной деградации мРНК и ее ковалентных посттранскрипционных модификаций.

3.3.1. Направленный транспорт, внутриклеточная локализация и депонирование мРНК

По завершении регулируемого синтеза РНК в процессе транскрипции она должна быть доставлена к месту трансляции, где сценарий координированной экспрессии генов получает свое дальнейшее развитие. При этом многие виды мРНК оказываются асимметрично распределенными в цитоплазме родительских клеток, пролиферация которых сопровождается избирательным попаданием (доставкой) мРНК в дочерние клетки. В итоге на ранних стадиях эмбриогенеза в зиготе и дочерних клетках происходит образование набора белков без соответствующей транскрипции их собственных генов, что обеспечивает их дальнейшую дифференцировку. Такой тип регуляции экспрессии генов через асимметричное распределение их РНК в цитоплазме играет исключительно важную роль в оогенезе животных, завершаемом

образованием зрелых яйцеклеток.

Структура и образование комплекса ооцит–питающие клетки у дрозофилы. Каждый из двух симметрично расположенных яичников дрозофилы содержит ~16 овариол, в передней части каждой из которых находится гермарий. В гермарии происходит асимметричное деление стволовых клеток, сопровождаемое воспроизводством стволовых и образованием коммитированных клеток, называемых цистобластами. Каждый цистобласт делится четыре раза с образованием 16 цистоцитов – группы клеток, связанных друг с другом цитоплазматическими мостиками, которые проходят через специализированные, связанные с цитоскелетом структуры, называемые кольцевыми каналами (ring canals). Лишь один из 16 цистоцитов становится ооцитом, остальные 15 – питающими клетками (nurse cells). Каждая из 16-клеточных зародышевых цист окружена соматическими фолликулярными клетками, что образует яйцевую камеру (1-я стадия развития). Задние части овариол составлены серией яйцевых камер, возраст которых уменьшается по мере удаления от оси тела насекомого. Наиболее удаленные от оси яйцевые камеры находятся в стадии развития 1, а ближайшие к оси – на конечной, 14-й, стадии развития. Для созревания яйцеклетки в яйцевой камере требуется три дня. В это время питающие клетки синтезируют большое количество РНК и белков, которые переносятся в созревающий ооцит. Большинство этих молекул бывают необходимы в первые два часа эмбрионального развития дрозофилы, до начала транскрипции в зиготе.

Выбор клетки, которая становится ооцитом среди 16 цистоцитов, не происходит случайно. Лишь две из 16 клеток связаны кольцевыми каналами с четырьмя другими, и одна из этих двух клеток всегда дифференцируется в ооцит. Большая цитоплазматическая структура, названная фьюсомой (fusome), в формировании которой участвует несколько белков цитоскелета, проходит через кольцевые каналы, объединяющие цистоциты, и участвует в детерминации ооцита. Единственный в 16-клеточном комплексе центр, организующий микротрубочки (microtubule organizing center – МТОС), расположен в проооците. МТОС объединяет нити микротрубочек, исходящих из всех 15 питающих клеток, которые проходят через кольцевые каналы. Поскольку в МТОС сходятся минус-концы микротрубочек, основанный на микротрубочках цитоскелет 16-клеточной конструкции структурно поляризован.

Такая направленность микротрубочек играет ключевую роль в упорядоченном транспорте РНК из питающих клеток в созревающий ооцит и специфическом распределении РНК в самом ооците, а следовательно, лежит в основе регулируемой экспрессии генов развивающегося организма.

Этот краткий экскурс в дифференцировку клеток зародышевой линии насекомого был предпринят, в частности для того, чтобы продемонстрировать, насколько самонадеянными являются попытки объяснения живого организма исключительно на молекулярном уровне. Если на заре молекулярной генетики (а именно эта часть дня сейчас за окнами лабораторий) такой подход оказывается плодотворным при изучении механизмов регуляции экспрессии генов у бактериофагов, то он явно не срабатывает в случае многоклеточных организмов. Смещение интересов современных генетиков от индивидуальных генов к организации целых геномов и геномике в силу большой сложности проблемы пока еще лишь намечается. И дело разворачивается медленно не только из-за сложности объекта, но скорее из-за очевидности альтернативного пути. Новых неизвестных генов, так же, как и новых видов насекомых, хватит еще не на одно поколение биологов. Однако для понимания организма как целого неизбежен переход к исследованию функционирования больших надмолекулярных комплексов, клеток и клеточных ансамблей.

Упорядоченное распределение РНК в ооцитах и ранних эмбрионах дрозофилы. Большинство РНК, которые позднее локализуются в растущем ооците, первоначально синтезируются во всех 16 клетках зародышевой линии. При этом временные различия аккумуляции РНК в ооцитах коррелируют с различиями в начале их синтеза, но не с различиями механизмов внутриклеточного и межклеточного транспорта.

Правильную локализацию РНК в ооците во многом обеспечивают микротрубочки его цитоскелета. Например, мутация в гене *homeless*, сопровождающаяся образованием ооцитов с двумя передними полюсами, приводит на стадии 7 к перемещению минус-концов микротрубочек к обоим полюсам, а их плюс-концов – в центр клетки. Это имеет следствием перераспределение РНК внутри мутантных ооцитов. РНК *bicoid* накапливается на обоих полюсах ооцитов, а РНК *oskar* – в центре клетки.

По мере созревания ооцита (стадии 7–8) могут происходить упорядоченные внутриклеточные перемещения импортированной РНК,

например между его полюсами. На заключительных стадиях созревания (10–14) питающие клетки усиливают синтез РНК и их перенос в ооциты на фоне перетекания ооплазмы, что сопровождается их специфическим внутриклеточным распределением.

Активация ооцитов после оплодотворения сопряжена с упорядоченными перестройками цитоскелета. Ядра зиготы дрозофилы претерпевают 13 синхронных делений без цитокинеза, образуя зародыш с несколькими тысячами ядер, окруженными общей цитоплазмой. Такой синцитий существует до конца 14-го клеточного цикла, в котором ~6000 ядер локализованы вблизи оболочки – кортекса. Затем впячивания мембраны формируют индивидуальные клетки бластодермы, в которой переднезадняя и дорсовентральная полярность положения клеток уже определена, и они содержат требуемые для дальнейшего развития материнские РНК. На стадии синцития и клеточной бластодермы начинается основной зиготный синтез РНК.

Механизмы регулируемой локализации РНК в клетках животных.

Механизмы, обеспечивающие специфическую локализацию РНК в клетках многоклеточных организмов, весьма разнообразны. Прежде всего, после завершения синтеза РНК имеет место ее направленная доставка из ядра в цитоплазму к месту хранения или трансляции. Однако часть внутриклеточных РНК не синтезируется в ядрах этих клеток, а экспортируется в цитоплазму из митохондрий или соседних клеток другой специализации. Кроме того, специальные механизмы контролируют направленное перемещение РНК в цитоплазме и ее избирательную деградацию.

Направленный ядерно-цитоплазматический транспорт РНК.

Одним из самых ранних посттранскрипционных событий, связанных с их направленной доставкой к месту внутриклеточной локализации и трансляции, является векторизованный экспорт РНК из ядра в цитоплазму. Несмотря на то что наличие механизма направленного переноса синтезированной РНК из одной части ядра к другой с последующим упорядоченным выходом в цитоплазму кажется очевидным, имеется мало экспериментальных свидетельств этому явлению. Большинство экспериментальных данных такого рода относится к ранним эмбрионам дрозофилы. В частности, направленный перенос продемонстрирован для транскриптов генов *hairy* и *fushi tarazu* в эмбрионах на стадии бластодермы, а также для мРНК гена *gurken*, которая на

стадии 8 ооцитов в их цитоплазме располагается по отношению к ядру строго упорядоченно. Помимо всего прочего, результаты этих экспериментов позволили предположить, что ядра как таковые обладают собственной полярностью, независимой от цитоплазматического цитоскелета.

Перенос РНК из клеток одного типа в клетки другого типа. Второй класс механизмов направленной доставки мРНК впервые обнаружен во время оогенеза дрозофилы. В этом случае РНК, синтезированная в питающих клетках, переносится в ооциты по кольцевым каналам. Питающие клетки связаны только с передними полюсами ооцитов, и часть поступающих из них РНК (например *bicoid*) специфически удерживается после их вхождения и остается в этой зоне ооцитов. У мутантов дрозофилы, для которых характерно взаимодействие питающих клеток с передним и задним полюсами ооцита, транскрипты гена *bicoid* локализованы на обоих полюсах ооцитов. В отличие от РНК, которые удерживаются на переднем полюсе ооцитов, другие РНК, преимущественно располагающиеся на заднем полюсе, активно переносятся внутри клетки к месту своей локализации. Такой перенос и внутриклеточная локализация транскриптов тесно связаны с цитоскелетом или обеспечиваются другими механизмами, например посредством их избирательной защиты от распада и деградации (подробнее см. ниже).

Импорт РНК из митохондрий. Цитоплазма заднего полюса ооцитов дрозофилы содержит большие, не связанные с клеточными мембранами органеллы – полярные гранулы, которые участвуют в образовании и дифференцировке клеток зародышевой линии. С полярными гранулами тесно ассоциированы митохондрии, осуществляющие экспорт в эти органеллы большой митохондриальной 16S рРНК, которая кодируется митохондриальным геномом. Функции этой РНК в ооците не ясны. Полагают, что она требуется для образования клеток переднего полюса.

Массовая деградация РНК, сопряженная с ее локальной защитой. Гетерогенное внутриклеточное распределение и депонирование РНК могут достигаться за счет ее избирательной защиты от деградации. Например, материнские транскрипты генов *nanos*, *cyclin B* и *Hsp83* концентрируются в плазме переднего полюса ранних эмбрионов дрозофилы, однако часть РНК остается нелокализованной. РНК переднего полюса переходит в отпочковывающиеся клетки во время дробления ооцита, тогда как остающаяся

РНК деградирует. Избирательную защиту РНК в плазме переднего полюса ооцитов обеспечивают вышеупомянутые полярные гранулы.

Другим примером избирательной дестабилизации и защиты мРНК является метаболизм материнской РНК гена *PCNA* в ооцитах асцидий. В цитоплазме этих ооцитов различают три основные зоны: центральную эктоплазму и периферийные участки, прилегающие к кортексу, между которыми находится миоплазма. Во время оогенеза не кодирующая белок РНК гена *YC* локализована вокруг ядра в перинуклеарном пространстве и постепенно переносится к кортексу. После оплодотворения эта РНК переходит в миоплазму и ассоциирует с цитоскелетом. 3'-Концевая последовательность *YC*-РНК комплементарна 3'UTR *PCNA*-РНК, а также 5'UTR РНК рибосомного белка *L5*. Вначале синтезируемая *PCNA*-РНК распределена равномерно в цитоплазме, однако постепенно миоплазма становится свободной от нее. Предполагают, что дестабилизация происходит в результате образования в миоплазме двухцепочечных *YC-PCNA* РНК-РНК-гибридов. В отличие от этого *L5*-РНК, несмотря на значительную гомологию с *YC*-РНК, не дестабилизируется, а накапливается в миоплазме. Истинные механизмы, лежащие в основе различий поведения этих двух транскриптов, остаются неизвестными.

Направленный транспорт РНК в цитоплазме. Многочисленные данные указывают на то, что направленный перенос РНК к специфическим внутриклеточным микрокомпартаментам осуществляется при участии и микротрубочек, и микрофиламентов цитоскелета эукариотических клеток. Такие выводы основаны, прежде всего, на результатах, полученных с использованием агентов, специфически нарушающих нормальное формирование частей цитоскелета, построенных из микротрубочек (колхицин, нокодазол, таксол) или микрофиламентов (цитохалазины). Информативными оказались и мутанты дрозофилы, и *Sacharomyces* с измененными компонентами цитоскелета.

Наиболее изученными в отношении механизмов транспорта являются транскрипты генов *bicoid* и *oskar* дрозофилы, первый из которых кодирует фактор транскрипции, а точные функции второго белка неизвестны. В процессе переноса от питающих клеток к месту своей окончательной локализации – переднему полюсу ооцита – РНК *bicoid* претерпевает несколько стадий ассоциации–диссоциации с компонентами цитоскелета. Внутриклеточный

перенос РНК в питающих клетках к кольцевым каналам осуществляется с участием минус-концевого мотора микротрубочек, так как подавляется их деполимеризующими агентами (колхицин, тубулозол С, нокодазол). Перемещение РНК через кольцевые каналы в ооцит не зависит от цитоскелета, поскольку не подавляется ингибиторами образования микротрубочек и микрофиламентов актина. Во время ранних этапов накопления транскриптов гена *bicoid* в ооцитах они снова ассоциируют с цитоскелетом, переходя в детергент-нерастворимую фракцию. На стадиях 8–11 развития ооцита РНК локализуются у его передней границы в неассоциированном с цитоскелетом состоянии, однако на стадии 14 транскрипты *bicoid* сильно компактизуются в передней части ооцита, находясь в связанном с микротрубочками состоянии. Наконец, в ранних эмбрионах эта РНК больше не ассоциирована с кортексом и цитоскелетом.

Зависимый от цитоскелета направленный перенос РНК происходит не только в оогенезе. Внутриклеточная локализация многих РНК в нейронах млекопитающих также происходит при участии микротрубочек. Функционирование того же механизма с участием микрофиламентов актина было обнаружено и при специфическом переносе ASH1-РНК в почкующиеся клетки дрожжей *S. cerevisiae*, а также при мембранной локализации мРНК β -актина в различных соматических клетках животных (фибробластах, миоцитах и т.п.).

Захват и фиксация РНК в сайтах внутриклеточной локализации.

На стадии 10В оогенеза дрозофилы происходит перестройка микротрубочек цитоскелета ооцита, которые формируют параллельные ряды вблизи его поверхности в субкортикальной зоне. Это индуцирует начало упорядоченного движения ооплазмы (*ooplasmic streaming*), в которое вовлекаются и многочисленные молекулы мРНК, поступающие в ооцит из питающих клеток. В данном случае РНК, которые специфически накапливаются в задней части ооцита, как бы отфильтровываются из потока ооплазмы путем захвата расположенными там специализированными структурами. Как следует из результатов генетического анализа, по крайней мере два белка Egalitarian и Bicardal-D вовлечены в этот процесс. В опытах с РНК *oskar* было установлено, что если первый белок требуется для первоначального захвата РНК *oskar*, то Bicardal-D не участвует в начальных этапах этого процесса, но необходим для

последующего удержания в задней части ооцита. Полагают, что в норме оба белка взаимодействуют друг с другом.

В захвате и удержании РНК на заднем полюсе ооцита дрозофилы также участвуют полярные гранулы – уже упоминавшиеся большие рибонуклеопротеиновые органеллы, не ассоциированные с клеточными мембранами. Доставляемая к ним в потоке ооплазмы РНК часто находится в составе больших транспортных частиц в комплексе с белками. В частности, РНК *bicoid*, накапливающаяся на переднем полюсе ооцита, в составе транспортных частиц находится в комплексе с белком Exuperantia, тогда как РНК *oskar* переносится к заднему полюсу связанной с белком Staufen.

Локализующие последовательности РНК и транс-действующие факторы, участвующие в направленной доставке РНК к местам ее внутриклеточной локализации. Использование гибридных РНК, содержащих в своем составе изучаемые и маркерные последовательности, является основным подходом к исследованию механизмов направленной доставки РНК к местам ее внутриклеточной локализации. В качестве маркерной последовательности часто применяют 5'-концевой фрагмент мРНК β -галактозидазы *E. coli*, которую обнаруживают в клетках *in situ* с помощью соответствующих антисмысловых гибридизационных зондов. Делеции, которые приводят к делокализации изучаемых РНК, также позволяют находить последовательности, направляющие РНК в определенные зоны цитоплазмы. В случае дрозофилы внутриклеточное поведение гибридных и мутантных РНК чаще всего исследуют у трансгенных мух. Для введения РНК в клетки также используют прямые инъекции и трансфекцию плазмид, экспрессирующих соответствующие гены.

Для идентификации *транс-действующих* факторов и их генов в настоящее время применяют три подхода. Первый из них, основанный на получении мутаций, нарушающих внутриклеточную локализацию РНК, используют для дрозофилы и *S. cerevisiae*. После создания мутантов осуществляют поиск и клонирование мутантных генов. При втором подходе исследуют факторы, специфически взаимодействующие с изучаемыми РНК. Наконец, обнаружение белков или РНК со специфической внутриклеточной локализацией также открывает путь к исследованию их генов.

Цис-действующие локализующие последовательности РНК. Все *цис-*

действующие локализирующие последовательности, обнаруженные до настоящего времени, расположены в 3'UTR их РНК. По аналогии с сигнальными аминокислотными последовательностями белков, наличие таких нуклеотидных последовательностей часто оказывается достаточным для направленной внутриклеточной доставки соответствующих РНК.

Создание одиночных локализирующих последовательностей может обеспечиваться альтернативным сплайсингом, механизм которого будет подробно рассмотрен в следующем разделе. Например, транскрипты гена *Cyclin B* дрозофилы процессируются по такому механизму. Короткая форма мРНК синтезируется преимущественно в раннем оогенезе. Она равномерно распределена в ооците до стадий 7–8. В отличие от этого более длинный транскрипт, содержащий дополнительную 3'-концевую последовательность длиной в 393 нуклеотида, в конце концов концентрируется в задней части ооцита и синтезируется в питающих клетках в позднем оогенезе (стадии 9–11). Эта же изоформа мРНК накапливается в перинуклеарном пространстве зародышей на стадии синцития.

В ряде случаев направленный транспорт мРНК (например *bicoid*) обеспечивается несколькими дискретными последовательностями, расположенными в 3'UTR. У этой мРНК локализирующий элемент BLE1 (*bicoid* localization element) длиной в 50 нуклеотидов является необходимым и достаточным для направленного транспорта мРНК, если он присутствует в виде двух копий. С такой двойной последовательностью взаимодействует белок Staufen, специфичный в отношении двухцепочечных РНК, который фиксирует РНК *bicoid* в передней зоне ооцита на поздних стадиях оогенеза. Этому способствуют и межмолекулярные взаимодействия РНК через их 3'UTR.

Кажущаяся функциональная избыточность характерна для регуляторных элементов ряда других РНК, содержащих несколько таких последовательностей, каждая из которых способна обеспечивать внутриклеточный транспорт РНК независимо от других. Примером такой РНК является транскрипт гена *orb* дрозофилы, который кодирует РНК-связывающий белок. Эта РНК начинает поступать в клетку на ранней стадии (1) развития ооцита и концентрируется сначала (стадии 2–7) на его переднем полюсе, а затем (стадии 8–10) на заднем. Такую направленную внутриклеточную доставку обеспечивает локализирующий регуляторный элемент длиной в 280 нуклеотидов.

Однако искусственное разделение данного элемента на два не инактивировало последовательность, хотя и снижало специфичность распределения РНК в цитоплазме.

Действие некоторых множественных локализирующих элементов мРНК оказывается аддитивным. Иными словами, разные сочетания таких последовательностей 3'UTR обеспечивают различную специфичность накопления РНК в цитоплазме. Примерами таких транскриптов являются уже рассматриваемая выше мРНК *bicoid*, а также мРНК Add-hts (изоформа N4), кодирующая белковый компонент цитоскелета дрозофилы. В последнем случае 3'UTR мРНК содержит центральную локализирующую последовательность ALE1 (Add-hts localization element 1) длиной 100–150 нуклеотидов, обеспечивающую накопление РНК вблизи поверхности ооцита на стадиях его созревания 7–8, и другую, необходимую для ее перемещения к передней части созревающего яйца на стадии 9, которая функционирует только в присутствии последовательности ALE1.

В ооцитах многие мРНК не транслируются до тех пор, пока они не доставлены к месту постоянной локализации. В частности, синтез белка, направляемый уже упоминавшейся мРНК *oskar*, регулируется по такому механизму. В норме трансляция этой мРНК начинается лишь после перемещения к заднему полюсу ооцита. Трансляция нелокализованной мРНК *oskar* сопровождается многочисленными дефектами развития зародыша насекомого. Контроль трансляции и локализация мРНК обеспечиваются разными последовательностями, расположенными в ее 3'UTR. В подавлении трансляции участвуют регуляторные элементы BRE (Bruno response elements), состоящие из трех дискретных сегментов А, В и С, взаимодействующих с белком-репрессором Bruno (молекулярная масса 80 кДа). Все эти сегменты обладают консервативной последовательностью длиной в 7–9 нуклеотидов: U(G/A)U(A/G)U(G/A)U, которые присутствуют в виде одной копии в сегментах А и В и представлены двумя копиями в сегменте С. Мутационные изменения этой последовательности нарушают взаимодействие с ними репрессора, а следовательно, и подавление трансляции мРНК *oskar* до завершения ее цитоплазматической локализации.

Транс-действующие белковые факторы. Уже упоминавшийся белок Staufen был обнаружен у мутантов дрозофилы с материнским типом

наследования дефектов локализации мРНК bicoid и oskar в цитоплазме ранних эмбрионов. Этот белок связывает двухцепочечную РНК, а также взаимодействует с одноцепочечным участком структуры типа "стебель–петля" в 3'UTR мРНК bicoid, обеспечивая ее правильную локализацию в ооплазме. Те же функции белок Staufen выполняет и в нейробластах, взаимодействуя в позднем эмбриогенезе с 3'UTR мРНК prospero, что необходимо для ее адекватной локализации в нейронах.

Белок Xsl дрозофилы с молекулярной массой 115 кДа взаимодействует с уже обсуждавшимися локализуемыми элементами BLE1, обеспечивая транспорт мРНК bicoid из питающих клеток и ее накопление в переднем полюсе ооцита. Этот белок выполняет свои функции как непосредственно связываясь с 3'UTR, так и обеспечивая роль посредника во взаимодействии белка Eupherantia с 3'UTR мРНК bicoid в локализующих РНП-частицах. Аллель eupherantia был получен вместе с аллелем staufer. Одновременное присутствие этих аллелей в геноме мутантных мух вначале обнаруживается по нарушению распределения РНК в питающих клетках. Белок Eupherantia накапливается в большом количестве вблизи оболочки передней части ооцита между стадиями 8 и 10. Его мутационные повреждения приводят к нарушению внутриклеточной локализации РНК bicoid в передней части ооцита и к прекращению ассоциации с микротрубочками цитоскелета. Поскольку в норме белок Eupherantia отсутствует на поздних стадиях оогенеза или в ранних эмбрионах, полагают, что он необходим для установления локализованного состояния РНК bicoid, но не для его последующего поддержания. Этот белок переносится в ооциты из питающих клеток в составе больших РНП-частиц через кольцевые каналы.

В настоящее время с помощью генетических и молекулярных методов обнаружено более десяти других *транс*-действующих белковых факторов, участвующих во внутриклеточной локализации мРНК дрозофилы, и несколько факторов, функционирующих в других организмах. Хотя многие из них оказались гомологичны известным РНК-связывающим белкам, прямое взаимодействие с локализуемыми РНК было продемонстрировано лишь для небольшой их части. Большинство РНК, обладающих такими свойствами, содержат несколько дискретных *цис*-действующих локализующих последовательностей, с которыми, как полагают, взаимодействуют

специфические *транс*-действующие факторы. Возможно, обнаруженные РНП-частицы, локализирующие РНК, содержат в своем составе такие факторы, а также ряд других белков, обеспечивающих взаимодействие РНК с системами ее цитоплазматического транспорта и фиксации в местах постоянной локализации. Большую роль в этих процессах отводят и комплементарным РНК-РНК-взаимодействиям.

Функции локализованных РНК. Направленная доставка и накопление РНК в определенных частях эукариотических клеток являются мощным механизмом, регулирующим экспрессию соответствующих генов в онтогенетическом развитии организмов. Прежде всего, локализованные РНК обеспечивают высокий уровень синтеза кодируемых ими белков в определенных частях клеток. Одновременно с этим предотвращается трансляция этих мРНК в других внутриклеточных компартментах как путем их избирательной деградации, так и с участием белковых репрессоров трансляции. Локализуемые изоформы мРНК с разной специфичностью внутриклеточной доставки, которые могут образовываться в результате альтернативного сплайсинга, обеспечивают синтез соответствующих изоформ белков в разных клетках или цитоплазматических доменах одной и той же клетки. При этом асимметричное внутриклеточное распределение мРНК, в свою очередь, может приводить к неравномерному их распределению в дочерних клетках, получающих специфические наборы мРНК. Это может быть одним из ключевых моментов последующей дифференцировки клеток-потомков. Ряд РНК, локализующихся в цитоплазматических компартментах, не кодируют белки. Во многих случаях их роль не ясна. Полагают, что они могут выполнять структурные функции в локализирующих РНП-частицах или органеллах, например зародышевых гранулах ооцитов. Кроме того эти РНК могут участвовать в фиксации локализуемых РНК в местах их накопления в цитоплазме.

От молекул ооцита дрозофилы к зародышу. Не все мРНК ооцита имеют своим источником питающие клетки. В частности, ген *gurken*, играющий ключевую роль в становлении переднезадней и дорсовентральной осей ооцита, транскрибируется в его ядре. Для этой РНК характерна строго выраженная внутриклеточная локализация. На стадии 7 она ассоциирована только с задним полюсом ооцита, на стадии 8 – с обоими полюсами, а на стадиях 8–10 – с

переднеспинной зоной. Белок *Gurken* является секретируемым TGF α -подобным фактором роста, который передает ключевой сигнал к окружающим ооцит фолликулярным клеткам, которые обеспечивают становление осей ооцита. Сначала сигнал в виде секретируемого фактора поступает к фолликулярным клеткам, окружающим заднюю часть ооцита, что обеспечивается соответствующей локализацией мРНК *gurken*. Это приводит к становлению переднезадней оси ооцита и поляризации его цитоскелета, построенного из микротрубочек, что является важнейшим условием дальнейшей локализации РНК. Сигнал в виде вышеупомянутого фактора роста, поступающий из переднеспинной части ооцита к питающим клеткам, приводит к становлению дорсовентральной оси яйцевой камеры. Локализация мРНК *gurken* в переднеспинной части ооцита, в свою очередь, является следствием зависимого от цитоскелета перемещения ядра ооцита в эту зону цитоплазмы.

мРНК *bicoid* начинает транслироваться в месте ее локализации на переднем полюсе раннего зародыша вскоре после оплодотворения. Поскольку на этой стадии зародыш представляет собой бесклеточный синцитий с многими ядрами, образующийся белок диффундирует к заднему полюсу зародыша, образуя градиент концентрации вдоль его переднезадней оси, которая максимальна в его передней части. Белок Bicoid является гомеодоменным фактором транскрипции. Он активирует синтез РНК соответствующих генов только в ядрах, расположенных в передней половине зародыша. При этом активация генов зависит от концентрации фактора. Например, ген *hunchback* содержит высоко аффинный сайт связывания белка Bicoid и активируется как в местах высокой, так и низкой концентраций фактора. В отличие от этого, такие гены как *orthodenticle* и *empty spiracles* обладают низкоаффинными сайтами связывания фактора и начинают транскрибироваться лишь в ядрах, расположенных ближе к переднему концу зародыша. По такому механизму формируются островки дифференцирующихся клеток передней части зародыша с разными группами экспрессирующихся генов. Однако это не все. Белок Bicoid обладает способностью функционировать не только как фактор транскрипции, но и как репрессор трансляции определенных РНК (например *caudal*), с 3'UTR которых он прямо взаимодействует. В результате передняя локализация мРНК *bicoid* приводит к образованию обратного градиента концентрации белка Caudal, РНК которого исходно не локализована. В итоге

пик концентрации этого белка имеет место в задней части зародыша, где он также определяет наборы транскрибирующихся генов.

Приведенные примеры охватывают далеко не все механизмы, детерминирующие дифференцировку клеток в развивающемся зародыше, и служат для иллюстрации ключевой роли посттранскрипционной внутриклеточной локализации мРНК в контроле экспрессии генов, транскриптами которых они являются. Как можно видеть уже из этих примеров, посттранскрипционное перемещение и метаболизм РНК являются важными элементами регуляции экспрессии генов в онтогенезе. Ниже некоторые другие аспекты метаболизма РНК как регулятора экспрессии генов будут рассмотрены более подробно.

3.3.2. Сплайсинг РНК в регуляции экспрессии генов

Разнообразные механизмы процессинга РНК в клетках были уже рассмотрены выше. Как оказалось, созревание мРНК играет важную роль и в регуляции экспрессии тех генов, транскриптами которых эти РНК являются. Одними из первых доказательства такого рода регуляции экспрессии генов были получены для морских ежей. Транскрипты, обнаруживаемые в цитоплазме клеток на стадии бластулы, можно видеть с помощью специфических ДНК-зондов в ядрах клеток кишечника. В ядрах клеток бластулы и дифференцированных клеток кишечника морских ежей синтезируются похожие наборы РНК, однако наборы цитоплазматических РНК, участвующих в трансляции, у этих типов клеток сильно различаются вследствие разного процессинга предшественников мРНК в ядрах.

Убедительные данные о контроле экспрессии генов на уровне процессинга РНК получены в экспериментах с вирусами клеток животных. Вирусы полиомы и SV40 не способны размножаться в недифференцированных стволовых клетках тератокарциномы. Однако, если стволовые клетки дифференцируются в клетки соматических типов, они становятся перmissive (чувствительными) к этим вирусам. Одной из причин исходной неперmissивности стволовых клеток к указанным вирусам является их неспособность осуществлять процессинг (сплайсинг) вирусных РНК в ядрах. Это сопровождается накоплением в ядрах стволовых клеток гигантских предшественников вирусных РНК, которые в конце концов разрушаются.

Данные такого рода в совокупности свидетельствуют о зависимости развития соматических клеток от образования в них специфических ферментов сплайсинга, которые отвечают за процессинг предшественников клеточных мРНК, необходимых для функционирования и дифференцировки соматических клеток.

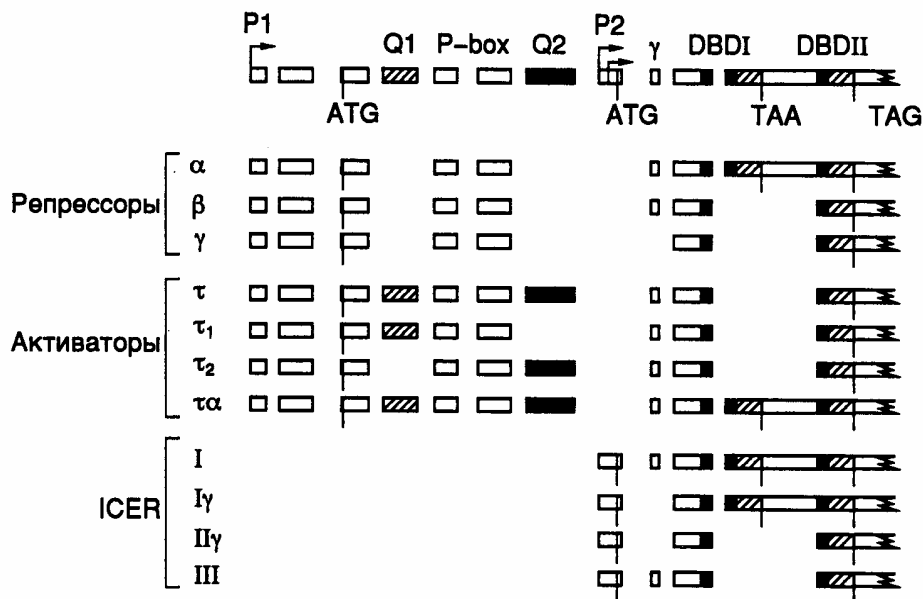


Рис. 1.35. Схема формирования активаторов и репрессоров транскрипции, кодируемых геном CREM и образующихся в результате альтернативного сплайсинга

В верхней части рисунка изображена схема гена CREM, включающая альтернативные промоторы P1 и P2, кодоны альтернативной инициации и терминации трансляции, а также экзоны, кодирующие различные домены факторов. Внизу представлены комбинации последовательностей экзонов, объединяющихся в зрелые мРНК, которые кодируют разных представителей групп репрессоров и активаторов транскрипции, а также ICER-белков

Альтернативный сплайсинг. Одним из примеров посттранскрипционных модификаций РНК, имеющих большое значение в регуляции экспрессии генов эукариот, является альтернативный сплайсинг. Сущность его заключается в том, что в результате посттранскрипционного процессинга предшественника мРНК, из которого в результате сплайсинга вырезаются некодирующие последовательности нуклеотидов, соответствующие интронам транскрибированного гена, образуются зрелые мРНК, различающиеся по своей первичной структуре. В результате в разных

клетках и тканях из одного и того же предшественника получают молекулы зрелых мРНК, которые объединяют в различных комбинациях последовательности экзонов транскрибированного гена.

На рис. 1.35 представлена схема экзонной структуры гена модулятора cAMP-респонсивного элемента человека (CREM), кодирующего целое семейство позитивно и негативно действующих изоформ фактора транскрипции, связывающихся с cAMP-респонсивными регуляторными элементами ДНК – CRE. Данное семейство факторов вовлечено в регуляцию экспрессии генов, участвующих в обеспечении эндокринных функций организма и сперматогенезе. Ген построен из функциональных модулей, заключающих в себе информацию двух промоторов, двух альтернативных ДНК-связывающих доменов, а также нескольких *транс*-действующих доменов этих факторов. ПремРНК транскрибируется с неиндуцируемого промотора P1 и подвергается альтернативному сплайсингу, специфичность которого определяется типом ткани, в которой он происходит. Например, во время сперматогенеза происходит переключение с синтеза репрессоров транскрипции CREM α , CREM β и CREM γ на образование CREM τ , который содержит в своей полипептидной цепи две дополнительные последовательности, обогащенные Gln (Q1 и Q2), необходимые для *транс*-активации транскрипции. Более короткий транскрипт, синтез которого иницируется на промоторе P2, процессируется с образованием зрелых мРНК, которые кодируют группу небольших молекул репрессоров (ICER). Инициация синтеза РНК на промоторе P2 активируется под действием ритмических гормональных сигналов эпифиза мозга.

Рассмотренный пример показывает важную роль альтернативного сплайсинга в расширении кодирующего потенциала генов эукариот, поскольку в результате использования такого механизма одна и та же последовательность нуклеотидов гена может кодировать несколько разных белков.

Комбинаторика объединения последовательностей нуклеотидов премРНК в процессе альтернативного сплайсинга может быть очень сложной, поскольку в него вовлекаются интроны и экзоны, а их объединение происходит с использованием различных точек сплайсинга, локализованных в этих последовательностях. Распознавание участков ДНК, в которых находятся точки сплайсинга, участвующие в конкретном акте посттранскрипционного

процессинга пре-мРНК, происходит с участием специфических белков в результате белково-нуклеинового взаимодействия.

Альтернативный сплайсинг у дрозофилы. Механизм регуляции альтернативного сплайсинга хорошо изучен у дрозофилы, использующей этот процесс, в частности при генетическом определении пола (рис. 1.36). Каскад реакций альтернативного сплайсинга в данном случае начинается с альтернативного выбора в пре-мРНК гена *sxl* последовательности экзона 3, содержащей кодон терминации трансляции. Если в результате альтернативного сплайсинга последовательность нуклеотидов, кодируемая экзоном 3, включается в зрелую мРНК самок, то при ее трансляции рибосомами образуется укороченная полипептидная цепь белка SXL, у которой отсутствует биологическая активность. Напротив, удаление такой последовательности в результате альтернативного сплайсинга приводит к образованию функционально активного белка, который, в свою очередь, участвует в регуляции альтернативного сплайсинга пре-мРНК гена *tra*, предотвращая взаимодействие мРНК U2AF⁶⁵ с одним из 3'-концевых сайтов сплайсинга. Это обеспечивает соединение белка с другим менее эффективным сайтом сплайсинга, расположенным ниже первого, участвующего в определении пола.

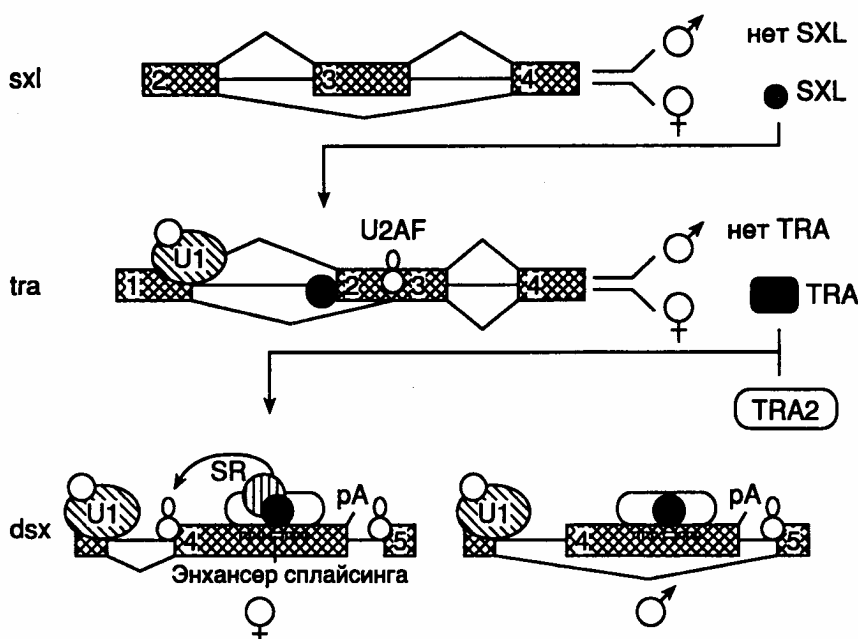


Рис. 1.36. Схема каскада реакций альтернативного сплайсинга, приводящих к определению пола у дрозофилы

Белок TRA стабилизирует связывание белка TRA2 с энхансером сплайсинга в *tra*-пре-мРНК. 1–5 – экзоны

Образующийся в результате белок TRA, содержащий аминокислотную последовательность, кодируемую экзоном 3 гена *tra*, в комплексе с другими белками TRA2 и SR-белком участвует в образовании мультипротеинового комплекса на специфической последовательности альтернативного экзона в мРНК гена *dsx*. Подобные регуляторные последовательности обнаружены также в мРНК позвоночных и вирусов животных и получили название энхансеров сплайсинга (они не имеют отношения к энхансерам, регулирующим транскрипцию эукариотических генов). В случае *dsx*-пре-мРНК образование такого энхансерного комплекса, включающего и белковые факторы, узнающие 3'-концевые сайты сплайсинга, способствует взаимодействию факторов сплайсинга с его низкоэффективным 3'-концевым сайтом, расположенным перед энхансером (и последовательностью экзона 4), а также вырезанию этого интрона. Образующийся в результате трансляции такой мРНК белок DSX блокирует экспрессию генов, обеспечивающих формирование у мух мужского фенотипа. В отсутствие TRA-белка в результате объединения экзонов 3 и 5 получается альтернативный белок DSX, который не препятствует экспрессии этих генов и подавляет дифференцировку самок.

Позитивная и негативная регуляция альтернативного сплайсинга в клетках животных. У животных описано много генов, пре-мРНК которых подвергаются альтернативному сплайсингу. Однако механизмы регуляции этого процесса не так ясны, как в только что рассмотренном примере с генами дрозофилы. Основную роль в выборе альтернативных сайтов сплайсинга у животных, по-видимому, играют общие факторы сплайсинга. В частности, установлено, что простое повышение концентрации SR-белков способствует предпочтительному выбору для сплайсинга проксимально расположенных, т.е. сближенных друг с другом, сайтов сплайсинга во многих пре-мРНК.

Механизм этого явления представлен на рис. 1.37 на примере гипотетической пре-мРНК, содержащей два альтернативных одинаково эффективных 5'-концевых сайта сплайсинга. Для простоты показано связывание с РНК только U1-мяРНП и SR-белка. В условиях недостатка U1-мяРНП (см. рис. 1.37,а) не все 5'-концевые сайты сплайсинга заполнены

мультибелковыми комплексами. Поскольку наиболее эффективно в альтернативном сплайсинге используется внутренний сайт, максимально сближенный со своим 3'-концевым партнером, альтернативный 5'-концевой сайт вовлекается только в том случае, если ниже расположенный сайт не занят. Повышение концентрации U1-мяРНП приводит к тому, что оба 5'-концевых сайта сплайсинга взаимодействуют с белковыми комплексами и сплайсинг происходит преимущественно с участием его максимально сближенных сайтов (см. рис. 1.37,б). В этом случае SR-белки усиливают взаимодействие между расположенными по соседству U1-мяРНП-комплексами.

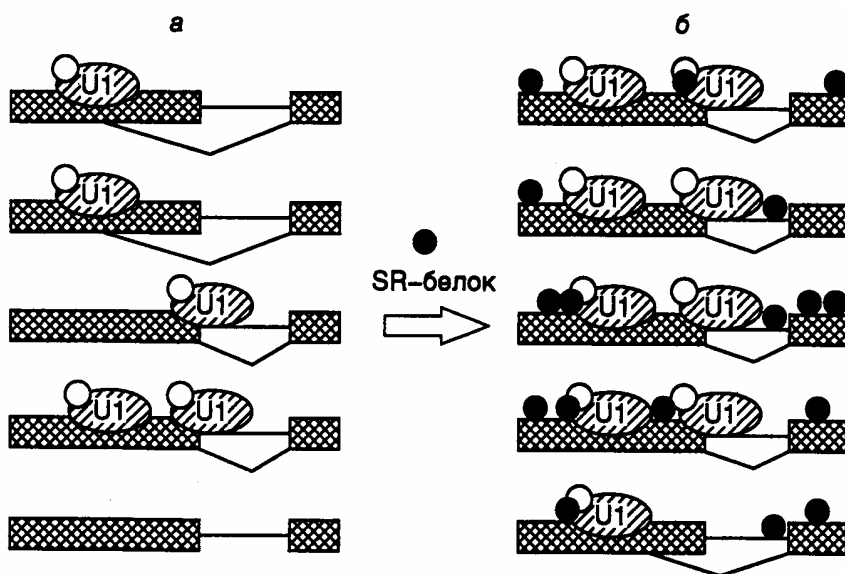


Рис. 1.37. Механизм действия белков SR при обеспечении выбора сплайсомой альтернативных сайтов сплайсинга

U1 – U1-мяРНП. Эффективность вырезания интронов: а – 40% проксимальных, 40% дистальных; б – 80% проксимальных, 20% дистальных

У многих генов животных и их вирусов обнаружены последовательности нуклеотидов, блокирующие сплайсинг. Некоторые из этих последовательностей способствуют формированию в пре-мРНК характерной вторичной структуры, которая маскирует сайты сплайсинга. Поскольку многие белковые факторы сплайсинга обладают ДНК-расплетающей активностью и вызывают локальное плавление спирализованных участков РНК, то дифференциальная экспрессия таких факторов оказывает влияние на выбор белками сплайсинга

определенных сайтов. Другие негативные регуляторные элементы альтернативного сплайсинга требуют для своего функционирования наличия *транс*-действующих факторов. Например, белок вируса иммунодефицита человека взаимодействует с последовательностью RRE пре-мРНК и, блокируя ее сплайсинг, обеспечивает экспорт непротранскрибированной пре-мРНК. Описаны экзоны, также оказывающие негативное влияние на сплайсинг. Кроме того, показано, что белки гетерогенных ядерных РНП A1, A1b, A2 и B1 выступают антагонистами SR-белков при взаимодействии последних с дистальными 5'-концевыми сайтами сплайсинга. Для индивидуальных белков этой группы (A и B) характерна субстратная специфичность, и они оказывают существенное влияние на выбор конкретных сайтов сплайсинга во время процессинга пре-мРНК, что в ряде случаев обеспечивается наличием высокоаффинных сайтов связывания для таких белков между конкурирующими 5'-концевыми сайтами сплайсинга. Например, наличие участка сильного связывания для белка A1, входящего в состав гЯРНП, между двумя 5'-концевыми сайтами сплайсинга приводит к преимущественному участию в процессинге дистального сайта сплайсинга (более удаленного от 3'-концевого сайта). Такое взаимодействие не оказывает влияния на связывание обоими 5'-концевыми сайтами сплайсинга U1-мЯРНП.

Рассмотренные механизмы позитивной и негативной регуляции альтернативного сплайсинга во многом гипотетичны. Однако существование внутриклеточных механизмов, с высокой точностью регулирующих этот процесс, в настоящее время доказано. Такая специфичность может достигаться за счет контролируемых изменений внутриядерной концентрации индивидуальных SR-белков и белков гетерогенных ядерных РНП-частиц, функционирующих в сочетании с регуляторными последовательностями нуклеотидов (энхансерами сплайсинга) и с сайтами сплайсинга определенной эффективности, локализованными в самих пре-мРНК. В некоторых случаях подобная регуляция требует участия дополнительных специфических факторов.

Переключение классов иммуноглобулинов. В процессе дифференцировки В-лимфоцитов происходит смена классов иммуноглобулинов, синтезируемых этими клетками. В начале дифференцировки образуются иммуноглобулины, обладающие способностью

как секретироваться в окружающую среду, так и интегрироваться в плазматическую мембрану В-лимфоцитов. Поверхностные иммуноглобулины выполняют здесь роль рецепторов антигенов и принадлежат к одному из двух классов белков: IgM или IgD. Белки этих двух классов обладают одинаковыми аминокислотными последовательностями в N-концевых частях, взаимодействующих с антигенными детерминантами, и различаются по C-концам. В начале процесса дифференцировки поверхностные иммуноглобулины представлены белками только класса IgM, а после первой стимуляции антигеном появляются белки обоих классов; далее, после завершения дифференцировки В-лимфоцитов, на поверхности клеток остаются иммуноглобулины только класса IgD. Именно они продолжают синтезироваться зрелыми В-лимфоцитами.

В основе феномена переключения классов иммуноглобулинов у В-лимфоцитов лежит механизм альтернативного сплайсинга. На любой стадии дифференцировки В-лимфоцитов синтезируются транскрипты генов иммуноглобулинов, в состав которых входят экзоны вариабельных частей иммуноглобулинов, а также их константных частей, как μ , так и δ , определяющих класс иммуноглобулинов (IgM или IgD соответственно). В результате альтернативного сплайсинга с последовательностями экзонов вариабельных частей в зрелой мРНК соединяются последовательности экзонов μ или δ , что и определяет, в конечном счете, класс иммуноглобулинов, образующихся в результате трансляции этих мРНК рибосомами. С помощью альтернативного сплайсинга В-лимфоциты решают вопрос и о том, какая из форм IgM – секретируемая или интегрирующаяся в мембраны – образуется в клетках. Так же как и в случае обсуждавшихся выше классов иммуноглобулинов, секретируемая и ассоциированная с мембранами формы IgM различаются по C-концевым аминокислотным последовательностям. Последовательности, кодирующие эти C-концевые домены иммуноглобулинов, добавляются в зрелые мРНК или удаляются из их предшественников в результате альтернативного сплайсинга.

Исследования последних лет показали, что регуляция экспрессии генов с помощью альтернативного сплайсинга широко распространена в клетках высших организмов. Использование такого изящного механизма позволяет одному и тому же гену кодировать несколько полипептидов, различающихся по

своим биологическим функциям, и контролировать их биосинтез во времени (на разных этапах дифференцировки клеток) и в пространстве (в различных органах и тканях). Все это повышает информационную емкость генома высших организмов без существенного увеличения его размеров.

Транс-сплайсинг. Из рассмотренных выше примеров следует, что во время постоянного или альтернативного сплайсинга интроны удаляются из предшественников мРНК в результате нескольких сложных реакций при участии крупного РНП-комплекса, называемого сплайсомой. Обычно 5'- и 3'-концевые сайты сплайсинга, участвующие в таком процессе, расположены на одной молекуле пре-мРНК, т.е. в *цис*-положении по отношению друг к другу. В отличие от этого в процессе *транс*-сплайсинга сплайсома выбирает для объединения 5'-сайт и 3'-сайт сплайсинга, которые располагаются на разных молекулах РНК. В настоящее время *транс*-сплайсинг обнаружен у простейших (трипаносом), евглен, круглых нематод и плоских червей, а также в митохондриях, хлоропластах и безрибосомных пластидах высших растений. Наиболее хорошо изучен механизм *транс*-сплайсинга у нематоды *Caenorhabditis elegans*, который подробнее будет рассмотрен ниже. Только у этих групп организмов (за исключением трипаносом, гены которых не содержат интронов) механизмы *цис*- и *транс*-сплайсинга функционируют в клетках на одних и тех же молекулах РНК.

Интересной особенностью транскрипции у нематод, сближающих их с трипаносомами, является то, что их РНК синтезируются в виде полицистронных предшественников, которые далее расщепляются до моноцистронных единиц, полиаденилируются и подвергаются *транс*-сплайсингу. Однако в отличие от трипаносом, у которых *транс*-сплайсинг ограничивается присоединением к акцепторному 3'-концевому сайту сплайсинга одной и той же лидерной РНК SL1 (spliced leader 1), нематоды обладают еще одним донорным сайтом сплайсинга, ассоциированным с SL2-РНК, молекулы которой присоединяются к внутренним сайтам полицистронных транскриптов. Таким образом, в клетках *C. elegans* сплайсинг осуществляется по трем механизмам (рис. 1.38). Это обычный *цис*- и *транс*-сплайсинг, в результате которых SL1-РНК присоединяется вблизи 5'-концов первичных транскриптов-предшественников зрелых мРНК, и, наконец, *транс*-сплайсинг с участием молекул SL2-РНК, которые объединяются с внутренними последовательностями нуклеотидов полицистронных

транскриптов. Все три типа сплайсинга проходят в основном с участием одних и тех же компонентов РНП-комплекса, представляющего собой сплайсому, которая использует одни и те же 3'- и 5'-концевые канонические последовательности сайтов сплайсинга. Подобно высшим организмам в состав сплайсомы нематод входят U1-U6-мРНК, из которых U2-U6 участвуют в *транс*-сплайсинге. Молекулы SL-РНК, содержащие донорные сайты сплайсинга, также входят в состав РНП-комплекса.

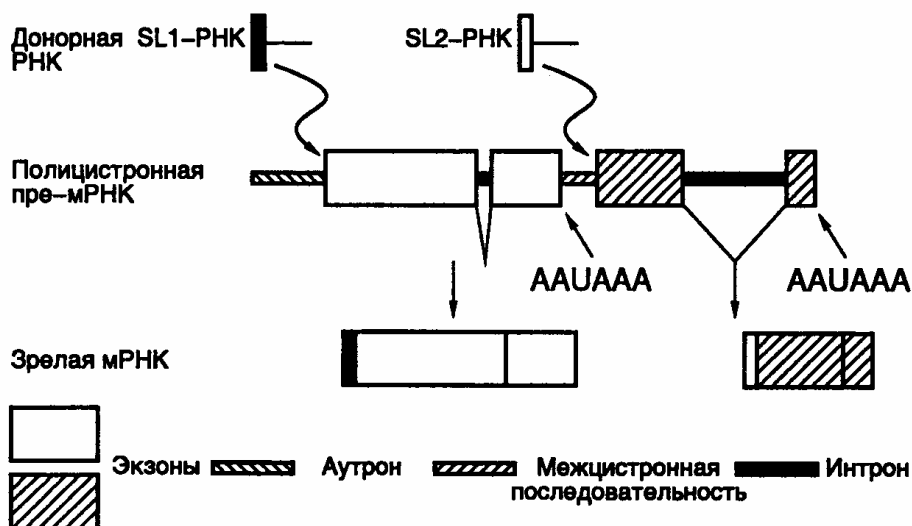


Рис. 1.38. Цис- и транс-сплайсинг ядерной полицистронной пре-мРНК у нематоды *C. elegans*

AAUAAA – сайты полиаденилирования

Последовательности нуклеотидов сайтов *транс*- и *цис*-сплайсинга идентичны и взаимозаменяемы в генно-инженерных экспериментах. Интроноподобные последовательности, расположенные на 5'-концах транскриптов, названные аутронами, подвергаются *транс*-сплайсингу, тогда как внутренние интроны вырезаются по механизму *цис*-сплайсинга. Единственным отличием гена, транскрипт которого подвергается *транс*-сплайсингу, является наличие на его 5'-конце последовательности аутрона, а не экзона. При перемещении интрона одного гена из внутреннего положения на 5'-конец другого или того же самого гена он начинает функционировать в качестве аутрона, т.е. подвергается *транс*-сплайсингу. Если размер такого аутрона превышает 50 п.о., то он нормально функционирует в *транс*-сплайсинге. При искусственном введении 5'-концевого сайта сплайсинга в аутрон на расстояние

более 50 п.о. от его сайта *транс*-сплайсинга заключенная между этими сайтами последовательность нуклеотидов в РНК подвергается обычному *цис*-сплайсингу. Таким образом, единственным сигналом, разрешающим *транс*-сплайсинг, является положение последовательности нуклеотидов вблизи 5'-конца пре-мРНК.

Зрелые, полностью процессированные РНК *C. elegans* содержат SL1- и SL2-РНК на 5'-концах, причем соблюдаются пропорции в содержании маркированных таким способом РНК внутри клеток. Следовательно, процесс *транс*-сплайсинга, в результате которого происходит выбор лидерной SL-РНК, каким-то образом регулируется. Предполагается, что белки, взаимодействующие с сайтами полиаденилирования соответствующих РНК, специфически взаимодействуют с белками мРНК, в состав которых входят SL-РНК, и тем самым осуществляют их выбор.

Функциональная роль *транс*-сплайсинга в жизненном цикле нематод неизвестна. Экспериментами *in vivo* продемонстрирована взаимозаменяемость генов, продукты которых подвергаются *цис*- или *транс*-сплайсингу. Известно, что сайты *транс*-сплайсинга часто располагаются непосредственно перед иницирующими AUG-кодонами мРНК. Это позволяет предположить, что SL-последовательности играют определенную роль в инициации трансляции. В частности, с помощью *транс*-сплайсинга могут удаляться лишние AUG-кодоны, находящиеся не в одной рамке считывания с основной кодирующей последовательностью нуклеотидов мРНК, или создаваться контексты нуклеотидов, которые обеспечивают эффективную инициацию синтеза белка рибосомами. Это, в свою очередь, может допускать определенную эволюционную гибкость в расположении соответствующих промоторов и создавать условия для возникновения новых сайтов инициации транскрипции на ДНК нематод.

3.3.3. Избирательная деградация мРНК

Время полужизни мРНК в клетках является важным фактором регуляции экспрессии генов. Феномен деградации мРНК как регуляторного явления впервые обнаружен у бактерий на заре развития молекулярной генетики. Уже тогда все РНК бактериальной клетки были разделены на два класса: стабильные и нестабильные. К первым до сих пор относятся рибосомные и

транспортные РНК, тогда как вторую, наиболее гетерогенную по составу группу образуют матричные РНК. Быстрая деградация мРНК после прекращения ее биосинтеза в результате регуляторных воздействий на транскрибируемые гены позволяет бактериальным клеткам легко адаптироваться к изменяющимся условиям окружающей среды и дает им ощутимые селективные преимущества перед клетками, которые не обладают таким регуляторным механизмом. Быстрая адаптация особенно актуальна для бактерий – одноклеточных организмов с коротким жизненным циклом.

Как было упомянуто, клетки эукариотических организмов способны выводить мРНК из трансляции, не разрушая ее. Это часто достигается регуляцией процессинга предшественников мРНК, образованием рибонуклеопротеидных комплексов – информосом, или специфической модификацией регуляторных последовательностей мРНК (см. ниже). Тем не менее, селективная деградация мРНК является распространенным механизмом регуляции экспрессии генов в клетках высших организмов. В эукариотических клетках время полужизни стабильных мРНК, таких как глобиновая мРНК, составляет ~17 ч, а время функционирования мРНК факторов роста не превышает 30 мин. Физиологические последствия таких различий очевидны. Если глобиновые мРНК продолжают транслироваться на протяжении всей жизни предшественников эритроцитов, то потребность в факторах роста ограничивается фазами клеточного цикла, непосредственно связанными с делением клеток. В частности, известно, что увеличение стабильности мРНК протоонкогена *c-fos* под действием мутаций сопровождается непрерывным делением мутантных клеток и образованием опухолей.

Время жизни эукариотических мРНК в цитоплазме чаще всего контролируется их 3'-концевыми нетранслируемыми последовательностями (UTR). Лабильные мРНК содержат в этой области одну или несколько AU-богатых последовательностей длиной около 50 нуклеотидов, называемых ARE-элементами (adenylate/uridylylate-rich elements), которые придают полирибонуклеотиду конформации, делающие его высокочувствительным к расщеплению нуклеазами. Искусственное введение таких последовательностей в 3'-концевые области стабильных мРНК приводит к их дестабилизации, что сопровождается резким уменьшением внутриклеточного времени полужизни гибридных мРНК. ARE-элементы были впервые обнаружены у генов цитокинов,

участвующих в воспалительных реакциях организма, и в настоящее время описаны для мРНК генов *c-fos*, *c-myc*, *nur77*, *junB*, β -интерферона, интерлейкина 3, гранулоцит-макрофагового колонийстимулирующего фактора (GM-CSF) и ряда других мРНК. Длина ARE-элементов варьирует от 50 до 150 нуклеотидов, они содержат несколько копий пентануклеотида AUUUA и много остатков U, перемежающихся остатками A.

Различают, по крайней мере, три класса ARE-элементов. Элементы 1-го класса (характерные, в частности для мРНК гена *c-fos*) содержат одну–три копии последовательности AUUUA, ассоциированные с U-богатой областью или гомополимером U. Для элементов 2-го класса (тип GM-CSF) характерно наличие двух копий перекрывающихся нонануклеотидных последовательностей UUAUUUA(U/A)(U/A), включенных в U-богатую последовательность. Наконец, элементы, принадлежащие к 3-му классу (тип *c-jun*), совсем не содержат последовательности AUUUA.

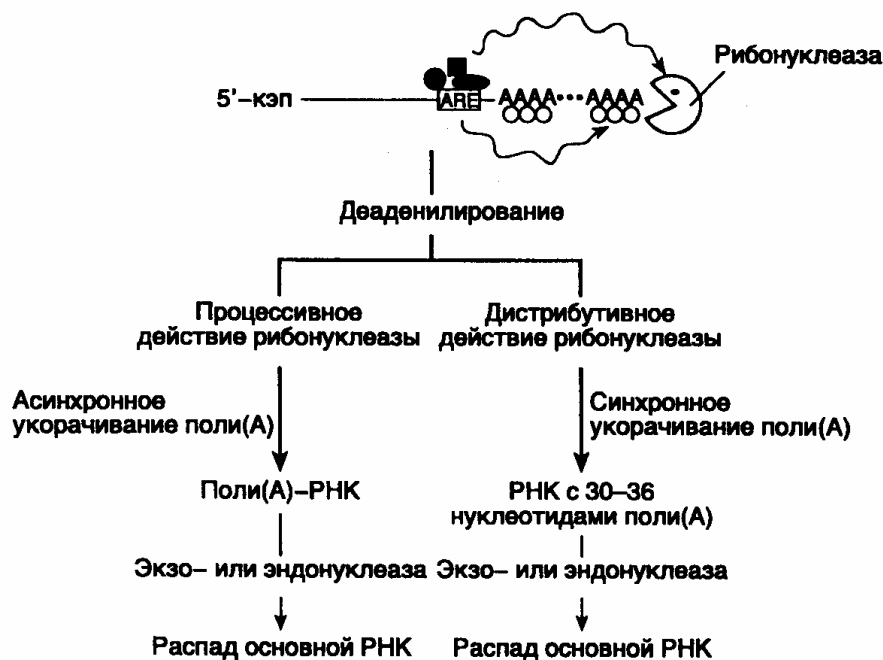


Рис. 1.39. Этапы ARE-зависимой деградации мРНК эукариот

Закрашенные геометрические фигуры – белки и факторы, специфически взаимодействующие с ARE-последовательностью. Их возможное влияние на деаденилирование показано волнистыми стрелками. Толщина стрелок указывает на интенсивность реакции

В соответствии с современной моделью ARE-зависимая деградация мРНК происходит двумя различными путями (рис. 1.39). Как уже упоминалось выше, этот процесс начинается с деаденилирования 3'-концевых поли(А)-последовательностей мРНК. Процесс деаденилирования проходит с разной скоростью в зависимости от класса ARE-последовательностей, присутствующих в мРНК. При наличии ARE-элементов 2-го класса (тип GM-CSF) деаденилирование отдельных молекул протекает асинхронно по процессивному механизму и заканчивается образованием молекул мРНК, не содержащих 3'-концевых поли(А)-последовательностей (поли(А)⁻-РНК). В отличие от этого ARE-элементы 1-го и 3-го классов направляют деаденилирование по дистрибутивному механизму. При этом отдельные молекулы деаденилируются синхронно, и на концах мРНК остаются поли(А)-последовательности длиной в 30–60 нуклеотидов. Скорость дальнейшей деградации процессированных мРНК может быть различной и зависит от регулируемой на уровне деградации мРНК экспрессии генов.

Таким образом, наличие в структуре мРНК различных классов специфических ARE-последовательностей, а также множественных белков-регуляторов, взаимодействующих с этими последовательностями, контролирует кинетику деградации мРНК в клетках. Аналогичные механизмы имеют место, в частности в селективной защите или дестабилизации мРНК в ответ на изменение внутриклеточных условий, например при гормональных воздействиях.

Изменение длины поли(А)-последовательностей мРНК. Выше было отмечено, что экпирование и полиаденилирование процессированных эукариотических мРНК повышают эффективность их трансляции рибосомами. В некоторых случаях эффективность трансляции мРНК, содержащих 3'-концевые поли(А)-последовательности, регулируется путем специфического изменения длины таких гомополимеров. В частности, у слизевиков *Dictyostelium* этот механизм играет ключевую роль в жизненном цикле. При переходе слизевиков от стадии вегетативного роста (амеба) к образованию плодового тела в клетках синтезируется новый набор мРНК. Одновременно резко укорачивается длина поли(А)-последовательностей мРНК, синтезированных и используемых в вегетативной стадии развития организма. В результате происходит переключение трансляции с мРНК, запасенных в

вегетативной фазе роста, на вновь синтезированные мРНК. Аналогичные эффекты наблюдаются в слюнных железах личинок дрозофилы, в процессе развития двустворчатого моллюска *Spisula*, а также в ооцитах шпорцевой лягушки *Xenopus*.

Хорошо изученным примером регуляции экспрессии генов на уровне полиаденилирования соответствующих мРНК является также дифференциальное полиаденилирование мРНК белка U1A – компонента U1-мРНК. Установлено, что 3'-концевые нетранслируемые участки (3'-UTRs) U1A-мРНК различных организмов содержат две копии консервативных последовательностей, которые обладают способностью связывать U1A-белок. В результате синтез избытка U1A-белка ингибирует полиаденилирование мРНК *in vivo*. При этом происходит подавление полимеризации АМР как следствие прямых контактов белка с поли(А)-полимеразой. После такой дестабилизации U1A-мРНК наблюдается пропорциональное снижение внутриклеточного уровня U1A-белка, т.е. имеет место регулирование его биосинтеза по принципу обратной связи.

Изменения кэп-группы мРНК. Своеобразный механизм контроля трансляции мРНК используется в онтогенезе некоторых бабочек. Например, у табачного бражника *Manduca sexta* эффективность трансляции определенных мРНК изменяется при ковалентной модификации их 5'-концевых кэп-групп. В ооцитах бражника запасенные мРНК содержат кэп-группу в виде неметилованного остатка гуанозина. Такие мРНК не транслируются рибосомами в бесклеточной белоксинтезирующей системе. Однако после оплодотворения кэп-группы мРНК быстро метилируются с образованием остатка 7-метилгуанозина и становятся активными матрицами при трансляции.

Этим примером мы завершим рассмотрение механизмов регуляции экспрессии генов путем посттранскрипционных модификаций мРНК и их предшественников. Как видно из изложенного материала, все регуляторные модификации структуры РНК оказывают влияние на экспрессию генов, транскриптами которых они являются, через изменение эффективности трансляции этих мРНК рибосомами. Фактически все вышеперечисленные механизмы регуляции экспрессии генов направлены на изменение матричной активности РНК-посредников (мРНК), с помощью которых происходит перенос генетической информации от транскрибируемых генов к белкам. Такие

механизмы оказывают регуляторное действие на уровень экспрессии регулируемых генов через изменение эффективности функционирования аппарата трансляции клеток организма. Тем не менее, среди механизмов, регулирующих экспрессию генов на уровне трансляции, многие непосредственно направлены на изменение функциональной активности самих компонентов аппарата трансляции: рибосом и многочисленных факторов трансляции, принимающих прямое участие в биосинтезе белков. Некоторые из этих механизмов будут рассмотрены в следующем разделе.

3.4. Регуляция экспрессии генов на уровне трансляции

В процесс биосинтеза белка рибосомами вовлекается большое количество мРНК, экипированных разнообразными регуляторными элементами. Даже в случае клеток дрожжей количество транслируемых видов мРНК превышает 6000. Регуляторные последовательности мРНК влияют на эффективность трансляции двумя основными путями: 1) изменением активности компонентов системы трансляции, взаимодействующих с регуляторными доменами мРНК; 2) изменением структуры регуляторных элементов самих мРНК. Как и в случае транскрипции, механизмы регуляции экспрессии генов на уровне трансляции осуществляют контроль эффективности всех основных этапов синтеза полипептидных цепей: инициации, элонгации и терминации.

3.4.1. Регуляция инициации трансляции

Инициация, т.е. сборка компонентов системы трансляции на 5'-конце мРНК, завершающаяся образованием первой пептидной связи, является важнейшей точкой приложения регуляторных воздействий на уровне трансляции. Эффективность инициации биосинтеза белка изменяется под действием различных гормонов, факторов роста и цитокинов, при изменении доступности питательных веществ и в условиях стрессовых состояний эукариотических клеток. Ключевую роль в этом играют факторы инициации трансляции eIF4E и eIF2.

Участие фактора инициации трансляции eIF4E в регуляции биосинтеза белка. Фактор eIF4E распознает кэп-структуры мРНК в составе

многокомпонентного фактора инициации eIF4F, что является необходимым этапом объединения мРНК с 40S субчастицей рибосом (подробнее см. раздел 2.5.2). Фактор eIF4E лимитирует инициацию трансляции. В большинстве клеток он присутствует в количестве 0,01–0,2 молекулы/рибосому, тогда как внутриклеточное содержание других факторов находится в пределах 0,5–3 молекулы/рибосому. Внутриклеточное содержание и активность фактора eIF4E регулируются на уровне транскрипции, посттрансляционно и путем взаимодействия с белками-репрессорами.

Регуляция биосинтеза eIF4E на уровне транскрипции. В ответ на действие сыворотки или факторов роста происходит многократное возрастание внутриклеточного содержания eIF4E-мРНК. Промотор гена этого фактора содержит два сайта связывания фактора транскрипции Мус, которые функционируют в искусственных гибридных генах. В соответствии с этим повышенный уровень экспрессии гена *c-тус* сопровождается возрастанием внутриклеточного содержания eIF4E-мРНК. Известно, что белок Мус участвует в регуляции пролиферации клеток. Поскольку фактор eIF4E сам по себе является ключевым регулятором роста и деления клеток, полагают, что его ген может быть одной из основных мишеней регуляторного воздействия белка Мус.

Регулируемое фосфорилирование фактора eIF4E. Фосфорилированное состояние полипептидной цепи фактора eIF4E коррелирует с повышенной скоростью трансляции. В митозе, характеризуемом низкой скоростью трансляции, уровень фосфорилирования eIF4E минимален. Количество фосфорилированных молекул фактора возрастает в ответ на внеклеточные воздействия гормонами, факторами роста, митогенами и цитокинами, а также в условиях повышенной нагрузки на сердце. У млекопитающих в ответ на все исследованные стимулы фосфорилирование полипептидной цепи происходит в основном в положении S209 (нумерация по полипептиду мышей). Остаток Thr в положении 210 фосфорилируется значительно реже. Поскольку в клетках, трансформированных онкогенами *ras* и *src*, наблюдается усиление фосфорилирования eIF4E, полагают, что в этом процессе участвуют MAP(ERK)-киназы (MAPK/ERK). Это участие может быть косвенным, так как в системах *in vitro* киназы ERK не обладают способностью фосфорилировать eIF4E. Недавно было показано, что общим субстратом протеинкиназ p38MAPK и ERK является протеинкиназа фактора eIF4E,

названная MNK1 (MAP kinase interacting kinase 1), которая в активном состоянии фосфорилирована. Поскольку MNK1 эффективно и специфически фосфорилирует фактор eIF4E *in vitro* по остатку Ser в положении 209, ее рассматривают в качестве основного кандидата, модифицирующего этот фактор и в живой клетке, после активации каскадов реакций с участием киназ ERK и p38 MAPK.

Семейство белков-репрессоров фактора eIF4E. Недавно были обнаружены небольшие белки (молекулярная масса ~12 кДа), названные 4E-BP1, 4E-BP2 и 4E-BP3 (eIF4E-binding proteins 1, 2 and 3), ингибирующие кэп-зависимую трансляцию после прямого взаимодействия с eIF4E. Образование комплекса eIF4E–4E-BP не изменяет сродство фактора к кэп-структуре, однако предотвращает его взаимодействие с eIF4G. Белки-ингибиторы 1, 2 и фактор eIF4G обладают гомологичной последовательностью аминокислот YXXXXLФ, где X – любая аминокислота, а Ф – алифатический аминокислотный остаток. Эта последовательность необходима для обсуждаемого белок-белкового взаимодействия.

Сродство ингибиторов семейства 4E-BP к фактору eIF4E регулируется через их фосфорилирование. Ингибитор 4E-BP1 был вначале идентифицирован как основной полипептид, фосфорилируемый под действием инсулина. Фосфорилирование полипептидных цепей ингибиторов предотвращает образование белок-белковых комплексов и происходит в присутствии гормонов (инсулин, ангиотензин, гастрин), факторов роста (EGF, PDGF, NGF, IGF1, IGFII), цитокинов (IL-3, GM-CSF), митогенов (TPA) и во время аденовирусной инфекции. В то же время в клетках некоторых типов тепловой шок и полиовирусная инфекция сопровождаются снижением уровней фосфорилирования ингибиторов. Все это указывает на прямое участие ингибиторов 4E-BP в регуляции трансляции у эукариот через взаимодействие с фактором eIF4E. По крайней мере, ингибиторы 4E-BP1 и 4E-BP2 являются субстратами протеинкиназы FRAP/mTOR – очень большого белка, принадлежащего к семейству киназ PIK, родственных киназам фосфатидилинозитола. Каскад реакций, завершающихся фосфорилированием этой киназы и, в конечном счете, белковых ингибиторов трансляции, запускается в ответ на вышеупомянутые внеклеточные стимулы киназой PIK3, фосфорилирующей OH-группу фосфоинозотида в положении 3.

Фактор eIF4E в регуляции роста и пролиферации клеток. Как следует из вышеизложенного, фактор eIF4E и его белковые ингибиторы являются специфическими мишенями протеинкиназ, активируемых в ответ на внеклеточные регуляторные воздействия. Это указывает на важную роль фактора в регуляции клеточного цикла. Действительно, микроинъекция eIF4E в покоящиеся фибробласты индуцирует в них синтез ДНК, а антисмысловые РНК к мРНК фактора резко увеличивают время прохождения клеток через G_1/S фазы клеточного цикла.

Сверхэкспрессия гена eIF4E приводит к характерным морфологическим изменениям в клетках HeLa и трансформирует иммортализованные клеточные линии грызунов. При этом происходит подавление апоптоза, индуцируемого в клетках истощением сыворотки. Кроме того, повышение внутриклеточного уровня фактора имеет место в опухолях различного происхождения. Все это делает фактор eIF4E объектом пристального внимания онкологов.

Фактор eIF2 как объект регуляторных воздействий. Как уже упоминалось выше, eIF2 представляет собой гетеротримерный белковый комплекс. Его α -субъединица фосфорилируется тремя известными киназами эукариот: у животных – HRI и PKR, а у дрожжей – GCN2. Фосфорилирование фактора предотвращает обмен GDP на GTP, опосредованный фактором eIF2B, и ингибирует трансляцию. Поскольку фосфорилированная форма eIF2 обладает повышенным сродством к eIF2B, последний становится эффективным конкурентным ингибитором формирования активного комплекса eIF2–GTP–Met-tRNA_i.

В качестве примера изменения эффективности трансляции мРНК через фосфорилирование фактора инициации eIF2 рассмотрим механизм контроля биосинтеза гемоглобина под действием гема. Этот пример интересен также и тем, что объясняет необходимость добавления гема в бесклеточные системы трансляции, получаемые на основе белков ретикулоцитов. Более подробно бесклеточные системы трансляции описаны в разделе 7.5.

Трансляция глобиновой мРНК в бесклеточной системе биосинтеза белка из ретикулоцитов кроликов в отсутствие гема (окисленной формы гема) сопровождается быстрым прекращением включения аминокислот в растущие полипептидные цепи, т.е. остановкой трансляции. Оказалось, что в отсутствие гема специфическая протеинкиназа фосфорилирует фактор инициации

трансляции eIF2, который в фосфорилированном состоянии прочно взаимодействует с другим фактором инициации eIF2B и в составе комплекса остается в связанном с рибосомами состоянии. В результате трансляция глобиновой мРНК останавливается. Гемин, находящийся в избытке в системе трансляции, взаимодействует с протеинкиназой и инактивирует ее. Протеинкиназа утрачивает способность фосфорилировать фактор eIF2 и, как следствие, блокировать трансляцию.

Координация синтеза глобинов на уровне трансляции происходит и в других случаях. Известно, что в диплоидной клетке человека имеются четыре активных α -глобиновых и лишь два экспрессирующихся β -глобиновых гена. Поскольку правильная сборка молекул гемоглобина предполагает участие эквимольных количеств полипептидных цепей α - и β -глобина, необходима координация биосинтеза этих белков, которая осуществляется на уровне инициации трансляции. Оказывается, α -глобиновая мРНК конкурирует с β -глобиновой мРНК за факторы инициации трансляции, однако β -глобиновая мРНК обладает большим сродством к факторам, что приводит к более высокой эффективности ее трансляции по сравнению с α -глобиновой мРНК. Предполагается, что в качестве фактора инициации трансляции, ответственного за предпочтительную трансляцию β -глобиновой мРНК, выступает кэп-связывающий белок.

Вышеописанные примеры показывают, как изменяется эффективность инициации трансляции определенных мРНК рибосомами при непосредственном воздействии на факторы инициации. Имеются и другие механизмы регуляции эффективности трансляции и, в конечном счете, регуляции экспрессии генов, реализующие свое действие через изменение эффективности инициации трансляции мРНК. Среди факторов, влияющих на эти механизмы, следует упомянуть, во-первых, разную эффективность ("силу") 5'-концевых областей инициации трансляции TIR (в частности последовательности Шайна–Дальгарно), необходимых для связывания рибосом в процессе образования инициаторного комплекса. Такие последовательности обеспечивают требуемую скорость трансляции соответствующих мРНК (подробнее см. раздел 7.2.6). Во-вторых, регуляция скорости инициации трансляции возможна за счет влияния пространственной

структуры 5'-концевого инициаторного района мРНК. Сворачивание этой части мРНК в стабильную пространственную структуру блокирует трансляцию. В-третьих, эффективная регуляция инициации трансляции определенных мРНК достигается за счет специфического взаимодействия инициаторных участков мРНК с белками-регуляторами, которые в данном случае выступают репрессорами инициации трансляции.

Белки, взаимодействующие с мРНК, как регуляторы трансляции.

Большинство регуляторных белков, взаимодействующих с 5'-концевыми TIR-последовательностями мРНК прокариот, являются негативными регуляторами трансляции. Классический пример такой регуляции экспрессии генов дают рибосомные белки *E. coli* – репрессоры собственного синтеза, которые предотвращают взаимодействие 30S субчастиц рибосом со своими мРНК. Оригинальный механизм репрессии использует рибосомный белок S15, который, взаимодействуя с TIR-последовательностью своей мРНК, стабилизирует предсуществующий псевдоузел. В результате SD-область мРНК становится ловушкой для 30S субчастицы рибосом, которая взаимодействует с ней, но не может иницировать синтез белка.

Аналогичные механизмы функционируют и у эукариот. В этом отношении хорошо изучена регуляция трансляции мРНК ферритина, синтазы δ -аминолевулиновой кислоты и субъединицы *b* сукцинатдегидрогеназы позвоночных животных. 5'UTR мРНК этих белков содержат регуляторный элемент IRE (iron-responsive element), с которым взаимодействует белок IRP (iron-regulatory protein), акцептирующий ионы железа. В отсутствие железа IRP связывается с IRE и блокирует трансляцию мРНК. Сродство IRP к IRE понижается в 50–100 раз, если он находится в комплексе с ионами железа. Этого оказывается достаточно для вовлечения соответствующих мРНК в трансляцию.

Цитоплазматические мРНК, не участвующие в синтезе белка в составе полисом, образуют нетранслируемые мРНП-комплексы. Кроме уже рассмотренных выше регуляторных белков, распознающих определенные последовательности мРНК конкретных видов, два белка обнаруживаются во всех мРНП в большом количестве: поли(А)-связывающий белок PABP (р70) и белок р50 с молекулярной массой ~50 кДа. Роль белка PABP в стабилизации мРНК и инициации трансляции уже обсуждалась. Теперь же целесообразно

рассмотреть регуляторные функции белка р50.

Белок р50, ассоциированный с цитоплазматическими мРНК-частицами. В отличие от белка РАВР, преимущественно ассоциированного с функционирующими полисомами, белок р50 является основным компонентом как неактивных мРНК, так и участвующих в синтезе белка. Белок р50 ретикулоцитов кроликов обнаруживает до 98% гомологии с факторами транскрипции животных, взаимодействующими с так называемым Y-боксом, *цис*-действующей регуляторной последовательностью ДНК СТАТТGGC/TC/ТАА. Факторы этого семейства преимущественно связывают одноцепочечную и апуринизированную ДНК, трехцепочечную Н-ДНК и РНК.

Отмечена двойственная роль белка р50 в регуляции трансляции: он может выступать как ингибитор и как активатор биосинтеза белка. При высоком отношении р50/мРНК (5–10 молекул белка на молекулу мРНК) имеет место ингибирование трансляции, при низком (до четырех молекул р50 на молекулу мРНК) – активация. Ингибирующая функция белка обнаружена при депонировании мРНК в ооцитах, а также в условиях сверхэкспрессии р50 в соматических клетках. Возможно, при высоких концентрациях белка происходит освобождение его С-концевых частей от контактов с РНК, приводящее к мультимеризации белка и переходу мРНК в конденсированное состояние.

Альтернативно, белок р50 выступает в качестве фактора трансляции в полисомах, активно синтезирующих белок. Полагают, что в этом случае он может облегчать инициацию трансляции, предотвращая неспецифическое взаимодействие мРНК с факторами трансляции, а также обеспечивая формирование у мРНК оптимальной пространственной структуры. Поскольку у р50 обнаружена РНК-расплетающая активность, он может способствовать сканированию 5'UTR мРНК прединициационным комплексом.

Антисмысловые РНК как регуляторы трансляции. Прокариотические антисмысловые РНК длиной 70–110 нт образуют структуры типа "стебель–петля", в которых стебель защищает эти РНК от деградации, а петля длиной шесть–восемь нт служит для первоначального взаимодействия с мРНК-мишенью. После образования комплексов РНК–РНК наблюдали изменение стабильности мРНК, эффективности процессинга РНК-мишени, терминации транскрипции или инициации их трансляции. Из этого видно, что антисмысловые РНК являются мощными природными модуляторами

экспрессии генов у прокариот. Данные о возможном участии природных антисмысловых РНК в регуляции трансляции у эукариот противоречивы.

Короткие ОРС в 5'-концевых лидерных последовательностях РНК как регуляторы трансляции. Около 10% мРНК растений содержат в своих 5'-концевых лидерных последовательностях более одного AUG-кодона. Некоторые из них удаляются с помощью альтернативного сплайсинга. Другие возникают в результате использования РНК-полимеразами альтернативных промоторов при инициации транскрипции соответствующих генов. Присутствие коротких ОРС в лидерных последовательностях мРНК, как правило, сопровождается снижением эффективности трансляции таких матриц. Функционирование этого механизма обнаружено в клетках млекопитающих, растений и дрожжей. Влияние коротких ОРС на трансляцию расположенных ниже последовательностей нуклеотидов мРНК недавно было детально исследовано с использованием искусственных генно-инженерных конструкций, в которых изменяли длину и число потенциальных сайтов инициации трансляции, предшествовавших генам-репортерам. Оказалось, что ингибирующее действие коротких ОРС возрастает с увеличением их длины. Даже одиночный AUG-кодон, снижает уровень трансляции ниже расположенных последовательностей, по крайней мере, в два раза. Короткие ОРС промежуточной длины (~30 кодонов) обладали пятикратным ингибирующим действием, а протяженные ОРС (>100 кодонов) полностью подавляли трансляцию следующих за ними последовательностей. Механизм ингибирующего действия коротких ОРС связан с тем, что они транслируются. Это снижает вероятность инициации трансляции на иницирующих кодонах, расположенных вслед за ними, поскольку процесс реинициации трансляции требует вхождения новых факторов инициации трансляции в инициаторный комплекс, включающий рибосому.

Трансактивация трансляции полицистронных РНК у вирусов. Предшественники геномной РНК вируса мозаики цветной капусты, а также их производные, подвергнутые альтернативному сплайсингу, являются полицистронными мРНК для многих вирусных белков. ОРС сближены друг с другом, и их не разделяют протяженные межцистронные последовательности. Такие РНК содержат внутренние AUG-кодоны, которые неэффективно используются для инициации трансляции в протопластах или трансгенных

растениях, однако начинают функционировать в присутствии вирусных генов-трансактиваторов (TAV) (рис. 1.40, д). В частности, трансактиваторная функция

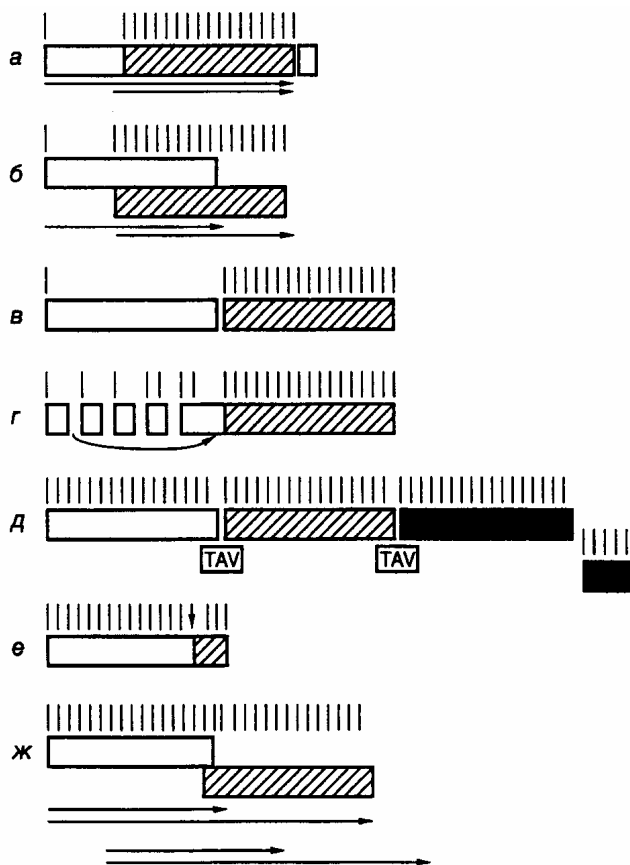


Рис. 1.40. Механизмы трансляции полицистронных мРНК у вирусов растений

Черные, серые и светлые прямоугольники изображают различные ОРС в полицистронных РНК, вертикальные линии над ними – AUG-кодоны, горизонтальные стрелки указывают направление перемещения рибосом во время шунта, вертикальная стрелка указывает положение частично супрессируемого стоп-кодона, TAV – гены-активаторы
 а, б, в – примеры ослабленного сканирования полицистронных мРНК, содержащих: а – два сайта инициации трансляции на одной ОРС (вирус мозаики вигны (cowpea)), б – перекрывающиеся ОРС (вирус желтой мозаики турнепса), в – две последовательно расположенные ОРС (вирус скрытой мозаики сливы); г – схема шунтирования сайтов инициации (вирус мозаики цветной капусты); д – трансактивация последовательно расположенных ОРС (тот же вирус); е – частичная супрессия стоп-кодона (вирус мозаики табака); ж – сдвиг рамки считывания (вирус желтой карликовости ячменя)

была показана для ОРС IV вируса CaMV и конкретных ОРС многих других

вирусов растений. Кодированный этим геном белок TAV специфически стимулирует трансляцию внутренних ОРС. В искусственных РНК трансактивация оказывается особенно эффективной, если в первой ОРС присутствует ~30 кодонов. Трансактивацию наблюдали для нескольких ОРС, которые были расположены ниже короткой первой рамки считывания. О том, что трансактивация происходит на уровне реинициации трансляции, свидетельствовали полярные эффекты вставок в такие полицистронные РНК последовательностей со вторичной структурой типа шпилек. Поскольку эффективность трансактивации зависела от длины первой ОРС, был сделан вывод, что трансактиватор действует прямо или косвенно на элонгирующие (или терминирующие) рибосомы. Оптимальная длина первой ОРС в 30 кодонов обеспечивает синтез пептида, длины которого достаточно для появления на поверхности транслирующей рибосомы. Предполагают, что на последующем этапе трансляции происходят структурные изменения рибосом, которые приводят к потере способности рибосом к реинициации трансляции и, как следствие, к трансактивации рибосом.

Функция трансактивации связана с центральной частью полипептидной цепи TAV. В процессе трансляции TAV взаимодействует с полисомами, а также с рибосомным белком (молекулярная масса 18 кДа) клеток дрожжей и растений. Трансгенные растения *Arabidopsis* и табака, экспрессирующие белок TAV, обладают ненормальным фенотипом и пониженной жизнеспособностью. Однако в настоящее время неясно, является ли это следствием способности белка обеспечивать трансактивацию трансляции или сопряжено с его другими, неизвестными активностями.

3.4.2. Регуляция элонгации синтеза полипептидных цепей

При обсуждении механизмов элонгации цепей РНК в процессе транскрипции была отмечена неравномерность прочитывания матричной ДНК РНК-полимеразами. То же самое наблюдается и во время элонгации растущих полипептидных цепей в процессе трансляции: не все участки мРНК транслируются с одинаковой скоростью. Прежде всего, рибосомы в процессе трансляции мРНК могут задерживаться на кодонах, соответствующих минорным изоакцепторным тРНК, присутствующим в клетке. В этом случае

внутриклеточная концентрация изоакцепторных тРНК лимитирует весь процесс трансляции. Кодоны, соответствующие минорным изоакцепторным тРНК, А.С. Спирин предлагает называть модулирующими, поскольку они могут изменять скорость трансляции соответствующих мРНК. Чем больше модулирующих кодонов в мРНК, тем медленнее она транслируется. В то же время клетка может изменять эффективность трансляции определенных мРНК путем адаптации внутриклеточных концентраций изоакцепторных тРНК к числу модулирующих кодонов этих мРНК. Было показано, в частности, что во время интенсивного синтеза фиброина в шелкоотделительных железах тутового шелкопряда внутриклеточный спектр изоакцепторных тРНК сильно меняется и становится идеально соответствующим потребностям белоксинтезирующего аппарата клеток, осуществляющего трансляцию мРНК фиброина.

Другим фактором, от которого зависит изменение скорости перемещения рибосомы вдоль транслируемой молекулы мРНК, является характерная пространственная структура матрицы. Для разворачивания индивидуальных участков пространственной структуры мРНК, обладающих неодинаковой стабильностью, требуется разное время, что находит отражение в различной скорости трансляции рибосомами индивидуальных мРНК.

Наконец, обнаружен ряд регуляторных белков, которые после взаимодействия с транслирующей рибосомой избирательно задерживают трансляцию в определенных местах мРНК. Например, у эукариот известна рибонуклеопротеидная частица, содержащая 7S-РНК, которая узнает особую N-концевую гидрофобную аминокислотную последовательность растущего полипептида, присоединяется к рибосомам и блокирует трансляцию до тех пор, пока рибосома не вступит во взаимодействие с мембраной эндоплазматического ретикула. Регуляция экспрессии генов на уровне элонгации трансляции широко распространена в живой природе. Во время многих вирусных инфекций скорость элонгации полипептидов зараженных клеток резко снижается. Это явление обнаружено, в частности у пикорнавирусов и вирусов осповакцины. Факторы элонгации трансляции могут быть мишенями различных регуляторных воздействий.

Запрограммированный сдвиг рамок считывания и неполная супрессия терминирующих кодонов во время элонгации полипептидных цепей. Процесс трансляции мРНК характеризуется высокой точностью, и даже

систематические "ошибки" трансляции могут быть генетически запрограммированными. У бактерий транслирующая рибосома может пропускать протяженные последовательности нуклеотидов мРНК, не прекращая синтеза единой полипептидной цепи. Такое явление неизвестно у эукариот. Однако у них в процессе декодирования кодона, находящегося в А-участке рибосомы, может происходить намеренное распознавание кодона "неправильной" аминоацил-тРНК или сдвиг рамки считывания у работающей рибосомы. Следствием этого бывает частичная супрессия терминации трансляции на терминирующих кодонах или синтез одной полипептидной цепи с использованием двух разных рамок считывания транслируемой РНК (см. рис. 1.40, е, ж). Хорошо изучены такие явления у ретровирусов и ретротранспозонов, которые используют сдвиг рамки считывания и супрессию терминирующего кодона для экспрессии гена *pol*, в результате которой синтезируется гибридный белок, N-конец которого является частью полипептидной цепи белка оболочки вируса (продукта гена *gag*). Тот же механизм используется и некоторыми другими вирусами животных, а также клетками дрожжей.

Сигналом к сдвигу рамки считывания у ретровирусов и ретротранспозонов служат гептануклеотидная последовательность типа X.XXY.YYZ (размечена в виде кодонов в рамке считывания 0), а также ниже расположенный регуляторный элемент, образующий определенную вторичную структуру в виде шпильки или псевдоузла. Предполагается, что сдвиг рамки происходит в тот момент, когда пептидил-тРНК, связанная с кодоном XXY в Р-участке рибосомы, и аминоацил-тРНК, взаимодействующая с кодоном YYZ в А-участке, одновременно сдвигаются на один нуклеотид назад и становятся напротив кодонов XXX и YYY транслируемой РНК. Сдвиг рамки считывания по этому механизму не всегда сопровождается образованием нового уникального белка, но часто приводит к синтезу небольшого числа вариантов полипептидных цепей, незначительно различающихся аминокислотными остатками в окрестностях сдвига рамки. Для осуществления сдвига рамки считывания РНК по такому механизму необходимо, чтобы обе молекулы тРНК в А- и Р-участках образовывали прочную связь с новыми кодонами, которые отличаются от первых только нуклеотидами в положении 3, допускающем неоднозначное соответствие антикодону. Процесс сдвига рамки вызывается

или усиливается структурным элементом РНК, перед которым работающая рибосома делает паузу в трансляции. Природа кодонов также важна для функционирования обсуждаемого механизма: из всех возможных ХХУ-кодонов в настоящее время в сайтах сдвига рамки считывания обнаружены только кодоны AAC, UUU, UUA и AAU. Терминирующие кодоны, часто обнаруживаемые сразу за сайтом сдвига рамки считывания, стимулируют сдвиг, так как вызывают остановку рибосомы. Эффективность сдвига рамки считывания может достигать 1–30%.

3.4.3. Регуляция терминации трансляции

Альтернативные сайты терминации трансляции могут быть использованы для расширения кодирующего потенциала определенных генов. Выше уже был рассмотрен пример, в котором в результате редактирования РНК в мРНК аполипопротеина В человека образуется новый терминирующий кодон, что приводит к синтезу в определенных тканях укороченного полипептида, кодируемого тем же самым геном, что и полипептид нормального размера.

Аналогичного эффекта система трансляции достигает посредством неполной терминации синтеза полипептидов на некоторых терминирующих кодонах. Из трех терминирующих кодонов наименее эффективным является UGA. Он чаще остальных ошибочно распознается транслирующей рибосомой как осмысленный (по-видимому, с участием триптофановой тРНК). В результате синтезируется более длинный полипептид, прекращение синтеза которого происходит на следующем терминирующем кодоне. В частности, такая ситуация наблюдается при трансляции РНК фага Q β . Цистрон белка оболочки фага заканчивается терминирующим кодоном UGA, который с небольшой частотой распознается рибосомами как осмысленный, что приводит к синтезу более длинного, чем белок оболочки, полипептида. Этот полипептид требуется для сборки полноценной (жизнеспособной) фаговой частицы и является жизненно важным для бактериофага Q β .

Для образования гибридного белка Gag-Pol ретровирусы типа С используют супрессию терминирующего кодона вместо сдвига рамки считывания. Супрессия происходит с эффективностью ~5% и сопровождается

ошибочным прочитыванием UAG-кодона глутаминил-тРНК. Терминирующий кодон UGA в том же положении декодируется как аргининовый, цистеиновый или триптофановый. Поскольку обычные терминирующие кодоны нормальных клеточных генов в этих условиях не супрессируются, делается вывод, что для осуществления супрессии терминирующий кодон должен находиться в определенном контексте. Запрограммированная супрессия терминирующих кодонов обнаружена, кроме того, у мРНК запасных белков растений, а также при трансляции геномной РНК вирусов растений. В последнем случае этот механизм используется для синтеза полипептидов РНК-зависимой РНК-полимеразы и удлинения белка оболочки.

Использование вышеописанных механизмов генетически запрограммированного сдвига рамки считывания транскрибируемой РНК или супрессии бессмысленных кодонов расширяет кодирующий потенциал геномов без физического увеличения их размеров. Еще более тонкий механизм изменения первоначальной генетической информации на уровне трансляции функционирует при введении в полипептидные цепи некоторых белков остатков селеноцистеина.

3.5. Синтез белков, содержащих остатки селеноцистеина

С помощью своеобразного механизма осуществляется передача генетической информации от генов к полипептидным цепям селенопротеинов с необычным аминокислотным остатком – селеноцистеином, входящим в их состав. У бактерий и млекопитающих известно более десяти ферментов, в состав активных центров которых входит остаток селеноцистеина, содержащего, в отличие от цистеина, атом селена вместо атома серы. Так, у *E. coli* гены форматдегидрогеназ H, N или O имеют в одной рамке считывания с кодирующей последовательностью нуклеотидов триплет TGA. Этому триплету в мРНК соответствует бессмысленный кодон UGA, на котором у подавляющего большинства других мРНК *E. coli* происходит терминация трансляции. Оказалось, что именно кодон UGA в мРНК вышеупомянутых генов кодирует селеноцистеин.

Встраивание этого аминокислотного остатка в полипептидные цепи регулируется весьма тонким механизмом. Перенос остатка селеноцистеина к рибосомам у *E. coli* осуществляется с помощью специальных молекул тРНК

(tRNA^{Sec}), которые на первом этапе соединяются с остатком L-Ser при участии серил- tRNA -синтетазы. Образовавшаяся серил- tRNA^{Sec} далее в результате многоступенчатого процесса под действием селеноцистеилсинтазы превращается в селеноцистеил- tRNA^{Sec} . Селеноцистеилсинтаза обладает высокой специфичностью и не взаимодействует с другими изоакцепторными серил- tRNA бактериальных клеток. Именно селеноцистеил- tRNA^{Sec} в процессе трансляции распознает в мРНК кодон UGA, но лишь в определенном контексте: для правильного узнавания UGA-кодона как осмысленного важна последовательность длиной в 45 нуклеотидов, расположенная вслед за UGA-кодоном. Кроме того, для правильного узнавания UGA-кодона селеноцистеил- tRNA^{Sec} необходимо участие белкового продукта гена *selB*, который является гомологом фактора элонгации трансляции EF-Tu и обладает высоким сродством именно к селеноцистеил- tRNA^{Sec} , но не к серил- tRNA^{Sec} . К тем же результатам, хотя и с использованием другого, не вполне понятного механизма, приводит встраивание в полипептидные цепи остатков селеноцистеина у млекопитающих.

Рассмотренный пример показывает, что при необходимости живой организм может изменять смысл стандартного генетического кода. В этом случае генетическая информация, заключенная в генах, кодируется более сложным образом. Смысл кодона определяется лишь в контексте с определенной протяженной последовательностью нуклеотидов и при участии нескольких высокоспецифических белковых факторов. Данный пример по-новому освещает понятие гена и смысл заключенной в нем генетической информации и не является единственным в своем роде.

Описано изменение смысла антикодона в tRNA путем посттранскрипционной модификации остатка цитозина с образованием так называемого лизидина. В этом случае происходит ферментативное присоединение Lys к гетероциклу цитидина в положении 2. В результате образовавшееся модифицированное основание – лизидин распознается как уридин, что изменяет специфичность антикодона модифицированной tRNA . Другое U-подобное азотистое основание – 5-карбамоилметилуридин (U^*), обнаружено в антикодоне tRNA^{Pro} (U^*GG), хотя в соответствующем гене этот антикодон детерминирован последовательностью CGG. По-видимому, здесь происходит посттранскрипционное дезаминирование цитозина с последующей

его гипермодификацией.

Таким образом, во всех приведенных примерах живым организмам недостаточно генетической информации, заключенной в их генах, для ее полноценной реализации в фенотипе. Пока не понятны причины, по которым организм избегает прямого кодирования соответствующих последовательностей нуклеотидов в своих генах, а предпочитает создание требуемых последовательностей в РНК путем посттранскрипционных модификаций первичных транскриптов. Такие факты меняют наше традиционное представление о генах как первичных носителях генетической информации.

3.6. Посттрансляционная регуляция экспрессии генов

Синтезом полноценного полипептида в результате трансляции кодирующей его мРНК рибосомами обычно завершается процесс передачи генетической информации от генов к белкам как у бактерий, так и у высших организмов. Однако в большинстве случаев при синтезе конечного белкового продукта эукариотическими клетками используются различные его модификации, в результате которых он и приобретает требуемые свойства.

Кроме того, необходимо иметь в виду, что конечный уровень содержания конкретных белков в клетке зависит не только от скорости их биосинтеза рибосомами, но и от скорости внутриклеточной деградации. Поэтому дифференциальная регуляция стабильности белков является важнейшим механизмом, регулирующим экспрессию генов у любого организма.

В этом разделе рассмотрены примеры посттрансляционных модификаций полипептидов, которые необходимы для получения физиологически активных пептидов или белков, т.е. завершения полноценной экспрессии конкретных генов.

3.6.1. Последствия фолдинга вновь синтезированных полипептидных цепей

В процессе трансляции растущие полипептидные цепи начинают приобретать высокоспецифическую пространственную структуру, которая формируется полностью вскоре после завершения их биосинтеза. Процесс сворачивания полипептидной цепи в правильную пространственную структуру

получил название *фолдинга*. В результате фолдинга в водных растворах у водорастворимого полипептида уменьшается свободная энергия, гидрофобные остатки аминокислот упаковываются преимущественно внутрь молекулы, а гидрофильные остатки располагаются на поверхности белковой глобулы. Гидрофобные области образуются и на внешней поверхности молекул белков, формируя полости активных центров, а также места контактов субъединиц мультимерных белков друг с другом и биологическими мембранами. Увеличение гидрофобности поверхности белков снижает их внутриклеточную стабильность, так как множество протеолитических ферментов гидролизуют с большой скоростью пептидные связи, образуемые гидрофобными аминокислотами или находящиеся вблизи от них.

Опасность протеолитической деградации для растущей полипептидной цепи возникает сразу после ее появления на поверхности транслирующей рибосомы. Установлено, что ~1/3 вновь синтезированных полипептидных цепей претерпевает протеолитический распад сразу же после завершения их синтеза рибосомами. Большинство вновь синтезированных белков избегает подобной участи благодаря образованию так называемого комплекса NAC (nascent polypeptide associated complex), ассоциированного с растущими полипептидными цепями. Имеется группа белков с молекулярными массами 21–33 кДа, которые взаимодействуют как с такой цепью, так и с самой рибосомой, предохраняя растущий полипептид от деградации путем формирования NAC. Когда же гидрофобная сигнальная последовательность синтезируемого белка достигает длины в ~70 аминокислотных остатков и покидает NAC, с ней взаимодействует комплекс белков SRP (signal recognition particle), который не только предохраняет гидрофобную часть растущего полипептида от ранней деградации протеиназами, но и направляет ее к месту назначения – к мембранам эндоплазматического ретикула.

Растущие полипептидные цепи, у которых отсутствует сигнальная последовательность, покидая NAC, взаимодействуют с обеспечивающей фолдинг системой, в состав которой входят, в частности шапероны Hsp70 и Hsp40. Эти белки теплового шока (Hsp – heat shock protein), образуя комплекс с растущей полипептидной цепью, предотвращают их неспецифическую агрегацию и деградацию под действием внутриклеточных протеиназ, способствуя их правильному фолдингу, происходящему с участием других

шаперонов. С другой стороны, Hsp70 принимает участие в АТФ-зависимом разворачивании полипептидных цепей, делая неполярные участки полипептидных цепей доступными действию протеолитических ферментов.

Различные сигнальные последовательности аминокислотных остатков обеспечивают направленную доставку вновь синтезированных белков к внутриклеточным органеллам и микрокомпартментам. Они же оказывают влияние на характер фолдинга, посттрансляционные модификации и метаболическую стабильность. Существуют, по крайней мере, пять биохимических процессов с участием вновь синтезируемых белков, контролируемых сигнальными последовательностями аминокислотных остатков. К ним относятся: транслокация белка через плоскость мембраны; внутриклеточный перенос белка без пересечения плоскости мембраны; химические модификации белка без гидролиза пептидных связей; расщепление некоторых или даже всех пептидных связей в белке; конформационные и иные пространственные изменения белков, включая фолдинг и олигомеризацию полипептидных цепей.

Время существования внутриклеточных белков может различаться на несколько порядков. Структурные и конститутивно экспрессирующиеся белки обычно обладают большой продолжительностью жизни. Напротив, регуляторные белки, как правило, быстро распадаются. Протеолитический гидролиз регуляторных белков, который может точно регулироваться, позволяет эукариотической клетке быстро переключаться с одной функциональной программы на другую. По времени полужизни белки животных разделяют на четыре группы: 1) очень быстро обновляющиеся белки (время полужизни – < 1 ч): белок-супрессор опухолей p53, продукты протоонкогенов *c-fos* и *c-myc*, орнитиндекарбоксилаза, циклины; 2) быстро обновляющиеся белки (время полужизни – 1–24 ч): тирозинаминотрансфераза, триптофан-2,3-диоксигеназа, γ -глутамилтрансфераза, Hsp70, РНК-полимераза I, рецептор инсулина, убиквитин; 3) медленно обновляющиеся белки (время полужизни – 1–5 дней): каталаза, калпаины, катепсины, протеасомы, тубулины, актины, альдолаза, лактатдегидрогеназа, аргиназа; 4) очень медленно обновляющиеся белки (время полужизни – >5 дней): митохондриальная фумараза, цитохромы b и c, миозин, гемоглобин, гистоны в интерфазном ядре, эластин, коллаген.

Большинство внутриклеточных белков заканчивают свое существование

в результате протеолитического гидролиза, превращаясь в небольшие пептиды и свободные аминокислоты, которые далее утилизируются в синтезе новых белков. Многие протеолитические ферменты используют в качестве субстратов индивидуальные белки, проявляя тем самым высокую специфичность. Тем не менее, в клетке имеется и множество протеиназ широкой субстратной специфичности, чья неразборчивость в субстратах компенсируется их строгой компартментализацией. Они локализованы в лизосомах и вакуолях, где гидролизуют любые белки после их попадания в эти органеллы. Такая компартментализация протеолитических ферментов является жизненно важным условием существования клетки. Система протеолитической деградации внутриклеточных белков с участием протеасом и убиквитина отличается от вышеописанных систем тем, что, обладая широкой субстратной специфичностью, она безопасна для окружающих белков и реагирует на регуляторные воздействия. Ниже будет рассмотрено функционирование некоторых из этих систем.

3.6.2. Специфические протеиназы в посттрансляционном процессинге белков

Одним из характерных примеров специфического действия протеиназ является активация предшественников (зимогенов) протеолитических ферментов (трипсина и химотрипсина) после их переноса от места синтеза (поджелудочная железа) к месту функционирования в пищеварительном тракте. Активация происходит вследствие протеолитического отщепления части полипептидной цепи зимогена, приводящего к появлению у укороченного полипептида требуемой ферментативной активности. Необходимость в посттрансляционной регуляции активности зимогенов очевидна – таким путем организм защищает клетки, синтезирующие протеолитические ферменты, от разрушения этими же ферментами.

Аналогичный механизм посттрансляционной активации используется во время биосинтеза инсулина и других пептидных гормонов. Предшественник инсулина синтезируется в виде длинного полипептида с тремя дисульфидными связями. Полипептид приобретает активность гормона только после протеолитического отщепления центральной части полипептида-предшественника и избыточной N-концевой аминокислотной последовательности. Две оставшиеся части предшественника, удерживаемые

дисульфидными связями, составляют активную молекулу инсулина.

В системах биосинтеза аденокортикотропного гормона (АКТГ) и эндорфина несколько небольших пептидных гормонов синтезируются в виде одного полипептида-предшественника. Дальнейшая судьба предшественника зависит от типа клеток, в которых он синтезируется. В клетках передней доли гипофиза этот полипептид расщепляется с образованием АКТГ, γ -липотропина и β -эндорфина, причем АКТГ является конечным продуктом протеолиза и может секретироваться для стимуляции синтеза стероидных гормонов в коре надпочечников. В клетках промежуточной доли гипофиза АКТГ расщепляется с образованием α -меланоцитстимулирующего гормона (α -МСГ).

3.6.3. Убиквитин-зависимая система протеолиза в регулируемой деградации белков

Убиквитин-зависимая система протеолиза проводит поиск потенциальной мишени для протеолитической деградации среди огромного числа внутриклеточных белков. Все белки несут в себе специфические сигналы деградации по аналогии с сигнальными последовательностями, которые направляют вновь синтезируемые белки к определенным органеллам или микрокомпартаментам клетки. Однако сигналы протеолитической деградации должны быть более сложными и разнообразными, так как с их помощью не только маркируются белки, удаляемые с помощью протеолиза, но и определяется время удаления, а также скорость их протеолитического расщепления. Для распознавания и декодирования таких сигналов в клетках эукариот имеется убиквитин-конъюгирующая система.

Как в ядре, так и в цитоплазме эта система отделена пространственно и функционально от протеолитических ферментов, организованных в протеасомы. Распознанные данной системой белки-субстраты маркируются путем ковалентного присоединения к ним молекул небольшого стабильного 76-звенного белка – убиквитина. В результате убиквитин соединяется С-концом с боковыми остатками лизина в субстрате. Наличие такой метки в белке, по-видимому, является первичным сигналом сортировки, направляющей образовавшиеся конъюгаты к протеасомам. В большинстве случаев к субстрату присоединяется несколько молекул убиквитина, которые организованы в виде

бусинок на нитке. Молекулы белков, содержащие убиквитин, по-видимому, являются для протеасом предпочтительными субстратами. Конъюгацию убиквитина с субстратом можно представить следующим образом (рис. 1.41). Вначале убиквитин-активирующий фермент (E1) связывает убиквитин, гидролизует АТФ и образует тиоэфирную связь между АМР и убиквитином с последующим переносом молекулы убиквитина на один из своих остатков Cys. Молекула активированного убиквитина далее соединяется с одним из ферментов семейства убиквитин-конъюгирующих ферментов (E2) и часто вслед за этим с убиквитин-лигазой (E3). Процесс конъюгации убиквитина с субстратом может катализироваться как самим E2, так и E2 совместно с E3. Белки E2 и E3 в клетках существуют в виде больших семейств, члены которых различаются по свойствам и внутриклеточной локализации. Мутации в генах семейства E2 у дрожжей показывают, что в ДНК-репарацию, прохождение клеточного цикла, биогенез пероксисом, а также в обеспечение устойчивости к тепловому шоку и ионам кадмия вовлечены разнообразные ферменты. Некоторые из ферментов E2 способны образовывать между собой гетеродимеры, которые, вероятно, в сочетании с различными белками E3 обеспечивают весь большой репертуар субстратных специфичностей убиквитин-конъюгирующих комплексов.

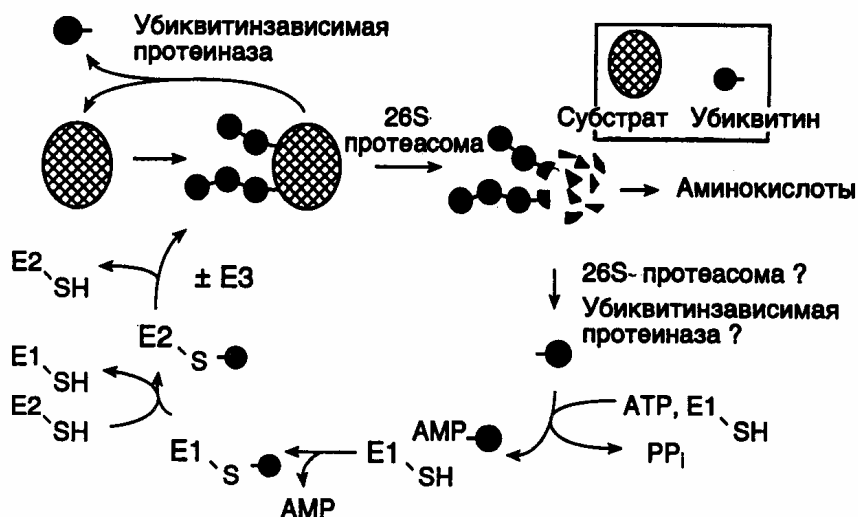


Рис. 1.41. Этапы функционирования убиквитинзависимой системы протеолиза

E1–E3 – ферменты, активирующие убиквитин с функционально-активными SH-группами, PP_i – неорганический пирофосфат

Структура и функционирование протеасомы. Белки, меченные цепями убиквитина, после взаимодействия с протеасомами расщепляются до коротких пептидов и свободных аминокислот в результате АТФ-зависимой реакции. У эукариот протеасомы присутствуют в цитозоле и ядрах, но не в других клеточных органеллах. Протеасомы выделяют в виде индивидуальных частиц с коэффициентами седиментации 20S и 26S. 20S частица является коровой частью 26S частицы, которая обладает протеолитической активностью. Коровая частица представляет собой белковый комплекс в виде цилиндра, через центральный канал которого "протягивается" молекула гидролизуемого белка. В настоящее время в качестве модельного объекта часто используют протеасомы археобактерии *Thermoplasma*, которые были закристаллизованы и исследованы с помощью рентгеноструктурного анализа.

Установлено, что протеасома *Thermoplasma* построена из двух (α и β) субъединиц. Кольцевые структуры на обоих концах цилиндра протеасомы составлены из α -, а центральная часть – из β -субъединиц в отношении $\alpha_7\beta_7\beta_7\alpha_7$ (рис. 1.42). Протеолитическая активность присуща β -субъединицам, причем их активные центры направлены внутрь полости цилиндра протеасомы. Кроме того, кольца α -субъединицы по обоим концам молекулы образуют узкие отверстия диаметром $\sim 13 \text{ \AA}$, через которые может пройти только развернутая цепь полипептида. Это механистически объясняет, каким образом протеасома избирательно расщепляет полипептидные цепи белков, меченные убиквитином. Прежде чем войти в контакт с активными центрами протеиназ, полипептидная цепь деградируемого белка должна быть развернута. Пептиды и аминокислоты, образующиеся в центральной части цилиндра протеасомы, покидают ее через переднее или заднее отверстия, сформированные α -субъединицами. Протеасомы *Thermoplasma* лишены специфичности в отношении деградируемых белков, и их функционирование требует наличия в N-концевой части β -полипептидов остатка Thr. Интересно, что остатки Thr эукариотических β -субъединиц являются мишенью для антибиотика лактацистина из *Streptomyces*, который ковалентно связывается с этими остатками, необратимо инактивируя протеасомы.

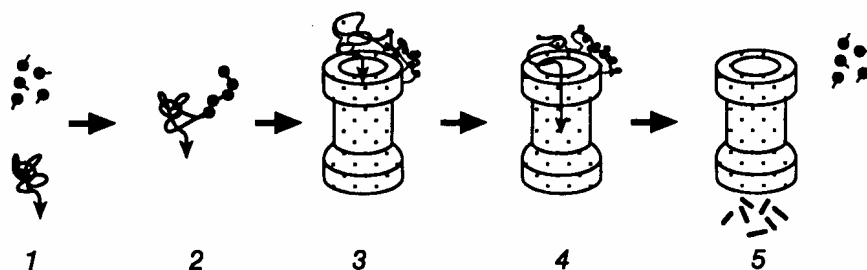


Рис. 1.42. Гипотетическая схема функционирования протеасомы убиквитинзависимой системы протеолиза

Молекулы убиквитина присоединяются к деградируемой полипептидной цепи (1,2) и конъюгат далее взаимодействует с 26S протеасомой (3). Узкий цилиндр изображает кóровую 20S субчастицу протеасомы, обладающую протеолитической активностью. Полипептидная цепь, разворачиваясь, входит в центральную полость 20S субчастицы, где последовательно подвергается протеолизу (4, 5). При этом цепи убиквитина отделяются от деградируемого белка

Локализация активных центров β -субъединиц внутри протеасомы затрудняет неконтролируемую деградацию окружающих белков. Отдельные субъединицы до включения их в состав зрелых протеасом синтезируются в виде неактивных предшественников, что предотвращает их преждевременное протеолитическое действие и является общим принципом биосинтеза протеолитических ферментов. Подобно тому как ферменты лизосом активируются только после их перемещения в соответствующий компартмент клетки, процессинг β -предшественников сопряжен с их включением в состав протеасом. Сборка протеасом начинается с образования гептамерных α -субчастиц (α_7), которые стимулируют аутокаталитическое удаление пропоследовательности предшественника β -субъединиц, что приводит к их упорядоченной самосборке с образованием гептамерных β -субчастиц (β_7), состыкованных с α -субчастицами. Две предварительно собранные половины протеасомы ($\alpha_7\beta_7$) далее ассоциируют друг с другом с образованием активных 20S протеасом. Очищенные 20S протеасомы эффективно расщепляют небольшие пептиды, но не способны гидролизовать интактные белки. Распознавание конъюгатов убиквитина и разворачивание белковой глобулы происходят с участием ~16 белков, ассоциированных с 20S протеасомой и образующих 26S комплекс. Эти белки способны объединяться в отдельные

комплексы, получившие название 19S-кэп-структур. 19S-комплексы ассоциируют друг с другом в присутствии ATP и с 20S протеасомами *in vitro*, присоединяясь к концам цилиндра.

Протеасомы обеспечивают не только полную деградацию полипептидов, но и участвуют в процессинге предшественников с образованием зрелых, активных белков. В частности, процессинг субъединицы p50 транскрипционного фактора NF-κB животных, сопровождающийся отщеплением и деградацией C-концевой части полипептида, происходит с помощью 26S протеасом.

Процессинг антигенов. В настоящее время сформировалось мнение о том, что протеасомы клеток животных участвуют в процессинге антигенов, представляемых молекулами первого класса главного комплекса гистосовместимости (МНС). Из семи β-субъединиц протеасом человека лишь три обладают протеолитической активностью. Эндогенный антивирусный агент γ-интерферон селективно стимулирует синтез неактивных β-субъединиц, одновременно подавляя образование активных. Две из этих стимулируемых субъединиц, названные LMP2 и LMP7, кодируются генами МНС. До конца не ясно, какой биологический смысл имеет замещение почти всех активных субъединиц на неактивные, однако предполагается, что такое явление может иметь отношение к процессингу антигенов, представляемых для распознавания молекулами МНС. Действительно, для мышей с дефектными LMP2 или LMP7 характерно нарушение презентации некоторых антигенов. По-видимому, пептиды, образующиеся в таких реконструированных протеасомах, по размерам или составу аминокислотных последовательностей соответствуют тому стандарту, который требуется для эффективного действия антигенных детерминант во время иммунного ответа.

3.6.4. Сплайсинг белков

Феномен сплайсинга белков, обнаруженный в 1990 г. в группой Т.Стивенса, пошатнул еще один постулат молекулярной биологии, в соответствии с которым последовательности нуклеотидов зрелых мРНК всегда колинеарны аминокислотным последовательностям кодируемых ими полипептидов. Во время сплайсинга белков удаление избыточной генетической информации из макромолекул происходит не на уровне пре-мРНК, как это

имеет место во время обычного сплайсинга, а на уровне синтезированного полипептида путем вырезания из его внутренней части короткой аминокислотной последовательности. Именно данный механизм отличает сплайсинг белков от повсеместно распространенного процессинга предшественников полипептидов, который, как известно, сопровождается только протеолитическим расщеплением полипептида-предшественника с образованием более коротких белков без изменения их внутренней первичной структуры.

Открытие сплайсинга белков было сделано при исследовании экспрессии гена *VMA1* дрожжей *S. cerevisiae*, известного также под названием *TFP1*, который кодирует субъединицу АТРаза вакуолей с молекулярной массой 119 кДа (рис. 1.43,а). При исследовании гомологии данного гена с генами других АТРаз микроорганизмов было установлено, что все они кодируют более короткие полипептиды с молекулярными массами ~70 кДа, и их гомология с геном дрожжей распространяется только на концевые последовательности нуклеотидов, резко нарушаясь в центральной части гена. Более того, использование инсерционного мутагенеза (генного нокаута) (см. раздел 9.1.3) для инактивации этого гена приводило к прекращению синтеза в клетках дрожжей полипептида с молекулярной массой 69, а не 119 кДа. Механизм такого явления прояснился после постановки двух контрольных экспериментов. В первом из них введение мутаций со сдвигом рамки считывания в сегмент гена *VMA1*, который кодировал центральную часть полипептида, отсутствующую в зрелом белке, приводило к прекращению синтеза всего белка рибосомами из-за возникновения новых терминирующих кодонов в этой рамке считывания мРНК. Подобное бы не происходило, если бы центральная часть пре-мРНК гена *VMA1* удалялась в результате сплайсинга. Во втором эксперименте исследование экстрактов клеток дрожжей с помощью антител к центральной части полипептида, отсутствующего в зрелой субъединице, обнаружило белковый продукт с предсказанной молекулярной массой 50 кДа. Он присутствовал в соотношении 1:1 к функционально активному полипептиду с молекулярной массой 69 кДа. Это указывало на то, что последовательности РНК центральной части гена транслируются рибосомами и соответствующая часть полипептида-предшественника удаляется посттрансляционно. Дальнейшие исследования данного явления полностью подтвердили

предположение о том, что внутренняя аминокислотная последовательность из предшественника с молекулярной массой 119 кДа отщепляется в результате сплайсинга на уровне его полипептидной цепи (см. рис. 1.43,а). Внутренняя часть полипептида, удаляемая в результате белкового сплайсинга, получила название интеина, а наружные N- и C-концевые части – экстеинов. Позднее белки, изменяемые посттрансляционно в результате белкового сплайсинга, были обнаружены у многих микроорганизмов (см. рис. 1.43,б). Во всех случаях аминокислотные последовательности интеинов фланкированы короткими консервативными последовательностями. На основании особенностей первичной структуры интеинов и экстеинов разработаны алгоритмы поиска белков, подвергаемых сплайсингу посттрансляционно, в соответствующих базах данных. С использованием этого алгоритма, в частности обнаружена последовательность интеина в продукте гена *dnaB* хлоропластов красной водоросли *Porphyra porphira*.

Для объяснения механизма белкового сплайсинга предложено несколько моделей. Одна из них, разработанная группой Ф. Перлера и включившая в себя многие черты ранних моделей, представлена на рис. 1.44. В соответствии с этой моделью белковый сплайсинг является аутокаталитическим процессом и для своего осуществления не требует других кофакторов в дополнение к самому полипептиду-предшественнику. Событием, запускающим белковый сплайсинг, может быть: а) нуклеофильная атака ОН-группой консервативного остатка Ser, расположенного в C-концевом сайте сплайсинга, по карбонильной группе пептидной связи, расположенной в N-концевом сайте сплайсинга, или же: б) N–O-сдвиг в N-концевом сайте сплайсинга с последующей атакой ОН-группой C-концевого сайта и переэтерификацией. В результате функционирования обоих механизмов образуется разветвленное промежуточное соединение, которое при участии остатка Asp распадается на интеин-сукцинимид и экстеины, связанные друг с другом сложноэфирной связью. Эта связь преобразуется в пептидную в результате N–O-сдвига.

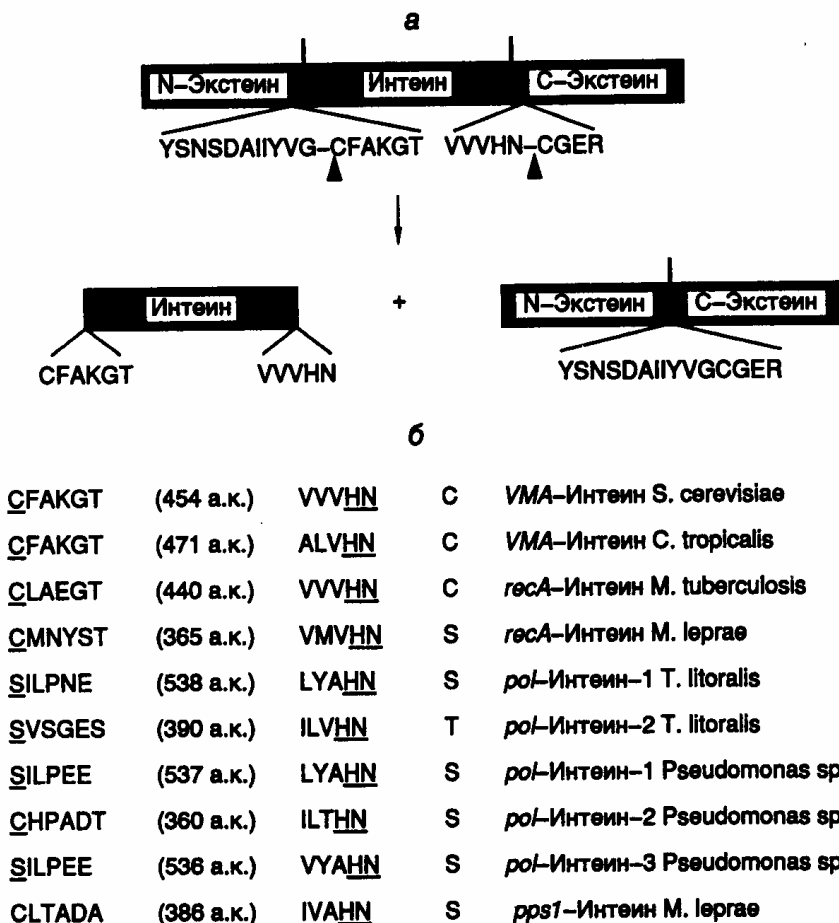


Рис. 1.43. Сплайсинг белка гена TFP1 *S. cerevisiae* (а) и перечень генов, белки которых подвергаются сплайсингу на уровне полипептидных цепей (б)

Стрелки указывают места разрыва полипептидной цепи белка-предшественника TFP1. Консервативные последовательности 10 известных интeинов в окрестностях сайтов сплайсинга изображены с использованием однобуквенного кода

Генетическая мобильность последовательностей интeинов в ДНК является одной из самых больших неожиданностей, обнаруженных после их открытия. Оказалось, что полипептидная цепь интeина обладает эндонуклеазной активностью, которая обеспечивает транспозиции последовательностей интeинов в геноме. Аналогичный процесс, ранее получивший название "хоуминга интронов", описан у интронов группы I. В результате хоуминга происходит однонаправленный перенос копии последовательности ДНК интрона (или интeина) из гена, содержащего эту

последовательность, к аллельному гену, не содержащему ее. Данная цепь реакций инициируется эндонуклеазным разрывом обеих цепей ДНК в аллельном безинтеиновом гене, образуя под действием эндонуклеазы интеина. Далее с использованием последовательности ДНК интеина в качестве донора информации происходит репарация двухцепочечного разрыва, которая сопровождается конверсией гена с включением в его состав последовательности интеина. Это указывает на большое сходство механизмов хоуминга интронов группы I и интеинов.

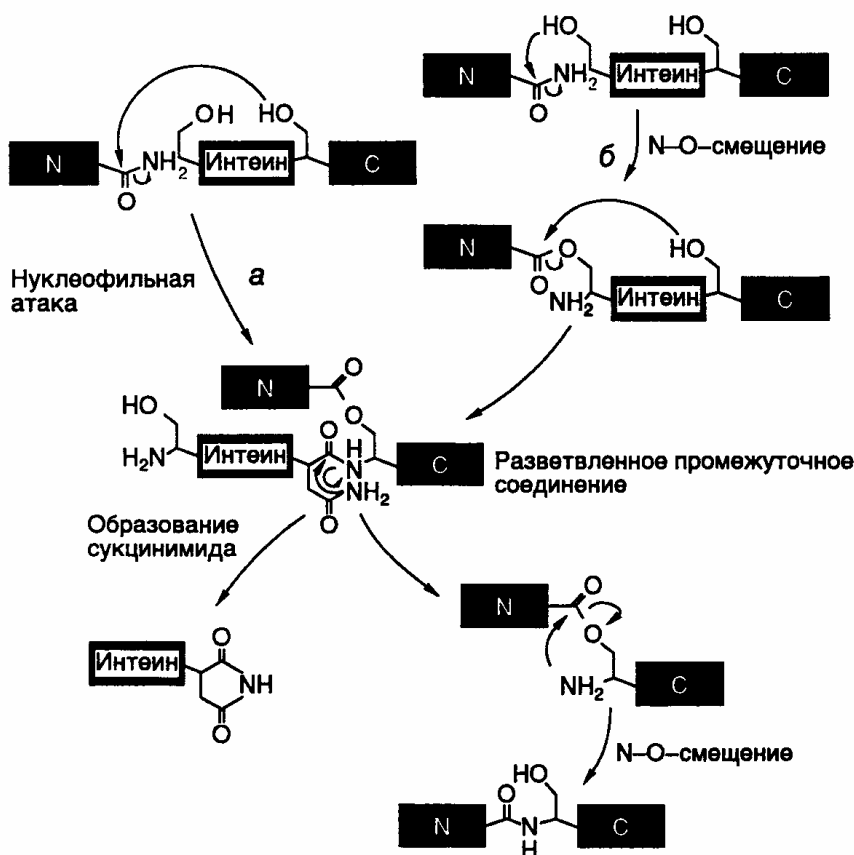


Рис. 1.44. Предполагаемый механизм сплайсинга белков

а, б – альтернативные механизмы инициации сплайсинга, N, C – N- и C-концевые экстеины белков-предшественников

Сплайсинг белков является сложной задачей для исследования *in vitro*. Процесс вырезания последовательности интеина происходит настолько быстро, что не удастся обнаружить промежуточные соединения. Проблему

удалось решить методами генной инженерии, вставив последовательность ДНК интеина из гена ДНК-полимеразы термофильной бактерии *Pyrococcus* в одну рамку считывания между геном белка, связывающего мальтозу, и частью гена парамизоина *Dirofilaria immitis*. Образующийся химерный белок не претерпевал сплайсинга при 12–20°, но при 37–65° сплайсинг индуцировался. Именно с помощью этого подхода удалось обнаружить разветвленные промежуточные соединения, образующиеся при сплайсинге белков. Кроме того, такой подход может стать продуктивным для биотехнологии, так как он позволяет нарабатывать токсические для клетки белки в виде химерных предшественников при низких температурах, которые затем простым повышением температуры быстро превращаются в полезные белковые продукты.

3.6.5. Другие посттрансляционные модификации белков

Многие белки и секретируемые пептиды претерпевают различные структурные изменения в результате котрансляционных и посттрансляционных модификаций, т.е. во время или после завершения их синтеза рибосомами. Подобные модификации существенно влияют на функциональную активность белков и пептидов, значительно расширяя возможности экспрессии генов, кодирующих эти молекулы, и их кодирующий потенциал. Одна из таких модификаций, а именно фосфорилирование факторов транскрипции протеинкиназами, рассмотрена выше. Кроме того, посттрансляционные модификации включают в себя гликозилирование остатков Asn в последовательностях Asn-X-[SerThr], N-концевое ацилирование, циклизацию N-концевого остатка Glu с образованием пироглутаминовой кислоты, C-концевое амидирование последовательностей освобождающихся пептидов, гидроксилирование остатков Lys и Pro, метилирование различных остатков аминокислот. Многие из перечисленных модификаций являются критическими для биологической активности пептидов. В частности, карбоксиамидирование C-концевого Gly активирует окситоцин и вазопрессин, а перенос сульфогруппы на остаток Tyr в холецистокиnine-8 оказывается критическим для проявления его активности в поджелудочной железе. N-Ацетилирование β-эндорфина блокирует его опиоидную активность, тогда как ацетилирование

меланоцитстимулирующего гормона усиливает его влияние на синтез меланинов. Поскольку большинство этих модификаций – тканеспецифические, пептиды, обладающие различной биологической активностью, должны быть доставлены к различным тканям в виде предшественников, где они претерпевают специфический процессинг.

Среди ковалентных посттрансляционных модификаций пептидов, синтезируемых рибосомами, занимает особое место эпимеризация *L*-аминокислотных остатков с образованием *D*-энантиомеров, присутствие которых оказывает принципиальное влияние на биологическую активность пептидов. Известно, что только *L*-стереоизомеры аминокислот участвуют в синтезе белка рибосомами. В природных белках *D*-аминокислоты обнаруживаются редко, как правило, в составе антибиотиков пептидной природы, которые синтезируются ферментативными комплексами микроорганизмов без привлечения рибосом. Другим источником *D*-аминокислот в белках может быть спонтанная рацемизация их *L*-стереоизомеров в составе полипептидных цепей в результате старения. Недавно обнаружено, что у ряда природных пептидов, обладающих биологической активностью, *D*-аминокислоты образуются во время посттрансляционных модификаций. В частности, это явление характерно для опиоидных пептидов, секретируемых кожей некоторых амфибий. Структура наиболее известных из таких пептидов – дерморфина и дермэнкефалина представлена на рис. 1.45,а. Активность этих пептидов как анальгетиков, по крайней мере, в 1000 раз превышает соответствующую активность морфина, и для ее проявления необходимо обязательное присутствие *D*-аминокислот в указанных положениях молекул. Дерморфиноподобные молекулы с теми же особенностями структуры были обнаружены в мозге крыс, а у амфибий найдена серия новых гексапептидов, делторфинов (Tyr-*D*-Ala-X-Val-Val-Gly, где X = Asp или Glu), которые являются агонистами δ -опиоидных рецепторов и для активности которых также необходимы *D*-аминокислоты. К пептидам той же природы относятся полициклические пептидные антибиотики грамположительных бактерий, в частности низин, субтилин и эпидермин, а также некоторые нейропептиды высших беспозвоночных. Механизм образования *D*-аминокислот в составе пептидов до конца не понятен, однако предполагается, что имеют место ферментативные реакции, в результате которых происходят последовательные

дегидрогенизация и гидрогенизация *L*-изомеров аминокислот (см. рис. 1.45,б). Поскольку в последнее время появились высокочувствительные аналитические методы, позволяющие обнаруживать следовые количества *D*-аминокислот в природных белках, можно предполагать скорое подтверждение гипотезы о том, что эта группа явлений имеет гораздо большее биологическое значение, чем ей сегодня отводится в системе биохимических превращений макромолекул.

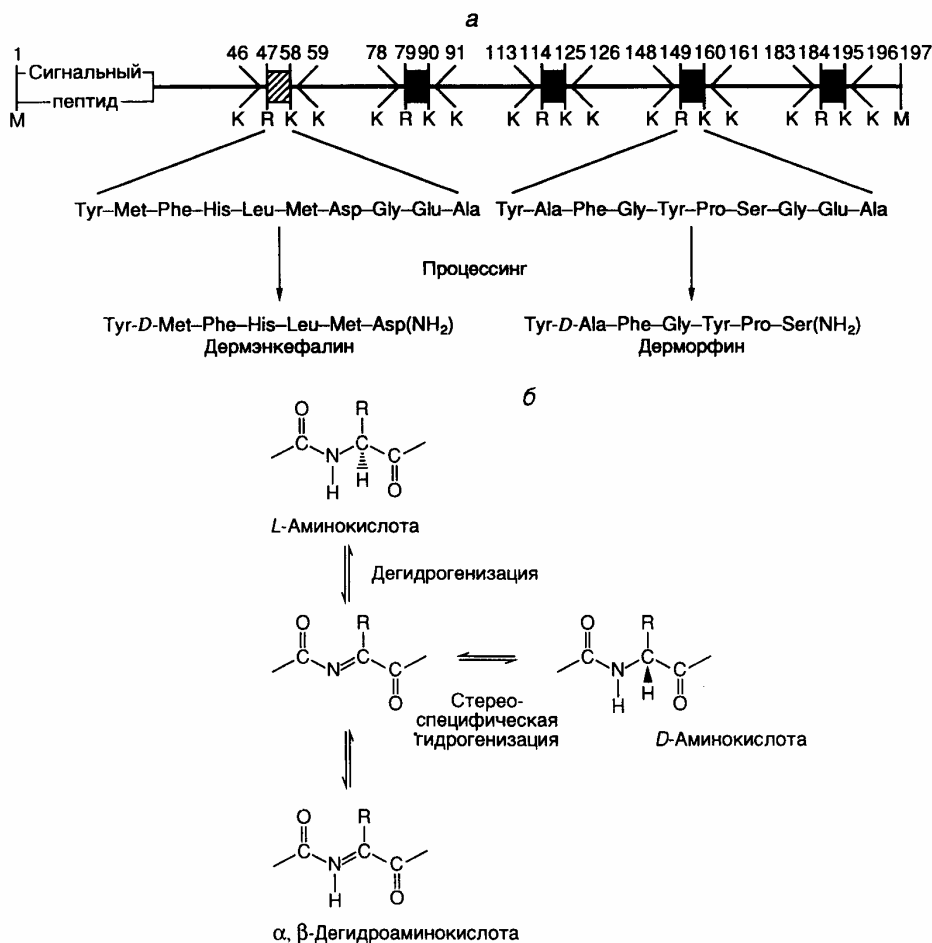


Рис. 1.45. Схема строения предшественника дермэнкефалина и дерморфина, а также их посттрансляционных модификаций с предполагаемым механизмом реакций

а – посттрансляционные модификации предшественника. Показано, что вслед за сигнальным пептидом в предшественнике располагаются пять высокоомологичных аминокислотных последовательностей, первая из которых является предшественником дермэнкефалина, а четыре других – дерморфина. Цифры обозначают положение аминокислотных остатков, приведенных ниже в однобуквенном коде; **б** – предполагаемый механизм реакции эпимеризации *L*-аминокислот в составе пептидов

Много этапов различных преобразований информационных макромолекул и белков отделяют генетическую информацию, заключенную в конкретных генах, от ее реализации в соответствующих фенотипических признаках клетки и организма, располагающих этой информацией. Такая реализация требует точной координации в функционировании многих генетических систем организма на всех основных уровнях экспрессии генов. Многоступенчатый и сложный характер регуляции функционирования генов, особенно у высших организмов, накладывает большие ограничения на возможности получения полноценной экспрессии рекомбинантных генов в гетерологичных системах. Тем не менее, расширение знаний об этих механизмах приближает момент, когда появится возможность целенаправленно изменять или корректировать работу отдельных генов организма, осуществлять наработку их полноценных белковых продуктов в искусственных генетических системах. О некоторых экспериментальных подходах к решению таких задач методами генной инженерии пойдет речь во второй части книги.

ГЛАВА 4. ВОСПРОИЗВЕДЕНИЕ ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ

Для того чтобы дочерние клетки по своей структуре и функциям были точной копией родительских клеток-предшественников, они должны получить от родительских клеток полный набор генетической информации в виде геномной ДНК, организованной в хромосомы, и внехромосомных генетических элементов (плазмид, митохондриальной и хлоропластной ДНК и т.п.). В связи с этим перед родительскими клетками встает задача создания точной копии генома и ее правильной передачи дочерним клеткам. Создание такой копии геномной ДНК в родительских клетках становится возможным благодаря наличию в них специальных ферментных систем, осуществляющих удвоение молекул ДНК. В результате реализации последовательности ферментативных реакций на матрице родительских ДНК происходит биосинтез дочерних молекул ДНК, которые являются точной копией исходных молекул. Этот процесс удвоения родительских молекул геномной ДНК во время воспроизводства клеток живого организма получил название *репликации*, или репликативного синтеза ДНК.

4.1. Репликация ДНК

Репликация ДНК происходит в соответствии с правилами Уотсона–Крика и наряду с биосинтезом РНК и белков является еще одним примером матричного синтеза биологических макромолекул. Во время репликации каждая из цепей родительской ДНК служит матрицей для синтеза комплементарной дочерней цепи. ДНК-зависимые ДНК-полимеразы осуществляют сополимеризацию низкомолекулярных предшественников ДНК – дезоксирибонуклеозидтрифосфатов (dATP, dGTP, dCTP и dTTP). Положение каждого последующего нуклеотида в строящейся цепи ДНК по правилам комплементарности однозначно определяется положением соответствующего нуклеотида матрицы.

При полимеризации дезоксирибонуклеозидтрифосфатов происходит освобождение молекул пирогосфата, который затем расщепляется неорганической пирогосфатазой, что делает реакцию полимеризации практически необратимой. Необходимо иметь в виду, что полимеризация нуклеотидов в процессе репликации происходит только в одном направлении:

от 5'-конца к 3'-концу строящейся цепи, и синтезированная молекула ДНК *антипараллельна* по отношению к ДНК-матрице. Репликация ДНК осуществляется по *полуконсервативному* механизму. Это означает, что одна из цепей дочерних молекул ДНК является частью родительской молекулы ДНК, а другая является вновь синтезированной.

Как и в случае биосинтеза других макромолекул клетки, процесс репликации условно разделяют на три основных этапа: инициацию, элонгацию и терминацию. Чтобы молекулы ДНК-полимераз могли начать синтез ДНК, им необходима *затравка* (праймер) – короткий олигодезоксирибонуклеотид или олигорибонуклеотид, комплементарный соответствующему участку ДНК-матрицы, у которого на конце имеется свободная 3'-ОН-группа. Неспособность молекул ДНК-полимераз самостоятельно без затравки начинать синтез ДНК принципиально отличает эти ферменты от других ферментов матричного синтеза – РНК-полимераз.

В соответствии с моделью Жакоба и соавторов (1963 г.) *репликоном* называют молекулу ДНК, способную к автономной репликации. Репликон содержит все необходимые гены и регуляторные последовательности, которые обеспечивают регулируемое удвоение его ДНК. Участок репликона, в котором начинается репликация, получил название *репликатора* или *области начала репликации* (replication origin) (у *E. coli* – *oriC*). При инициации репликации *инициатор*, кодируемый репликоном (у *E. coli* – белок DnaA), взаимодействует с репликатором.

4.1.1. Белки, участвующие в репликации ДНК

Процесс репликации ДНК осуществляется с участием множества белков, которые образуют сложный и эффективно работающий репликативный комплекс. В табл. I.16 представлены основные белки и ферменты, входящие в состав репликативного комплекса прокариотических и эукариотических организмов, и указаны их основные функции.

Таблица I.16

**Белки, входящие в состав репликативных комплексов прокариотических
и эукариотических организмов**

Белки в организмах			Функции компонентов комплексов
E. coli	Фаг T4	Вирус SV40 / человек	
DnaB	Белок 41	T-антиген	ДНК-хеликаза, стимулирует образование затравок на одноцепочечной ДНК
DnaC	Белок 59	»	Обеспечивает взаимодействие хеликазы и праймазы с ДНК, находящейся в комплексе с SSB-белком
SSB	Белок 32	RPA	Белок, связывающийся с одноцепочечной ДНК, стимулирует ДНК-полимеразы, облегчает вхождение хеликазы в репликативный комплекс
γ -Комплекс ($\gamma\delta\delta'\chi\psi$)	Белок 44/62	RFC	ДНК-зависимая АТРаза, обеспечивает связывание затравки с матрицей, стимулирует ДНК-полимеразу
τ - Белок	Белок 43 (?)		Обеспечивает сборку и димеризацию холофермента ДНК-полимеразы, необходим для образования инициационного комплекса

Таблица I.16 (окончание)

Белки в организмах			Функции компонентов комплексов
E. coli	Фаг T4	Вирус SV40 / человек	
β (β^*)-Белок	Белок 45 (?)	PCNA (?)	Стимулирует ДНК-полимеразу и ДНК-зависимую АТФазу, выполняет функцию "скользящего зажима", обеспечивающего процессивность репликации
Pol III ($\alpha\theta\epsilon$), минимальный фермент	Белок 43	Pol δ	ДНК-полимераза, 3'→5'-экзонуклеаза; α -субъединица Pol III катализирует полимеризацию, а ϵ -субъединица – является корректирующей экзонуклеазой
–	–	Pol ϵ	
–	–	Pol γ	ДНК-полимераза, осуществляет репликацию ДНК митохондрий, кодируется ядерным геном
DnaG	Белок 61	Праймаза, (Pol α)	Праймаза, синтез РНК-затравок
Лигаза	T4-лигаза	Лигаза I	Лигирование фрагментов ДНК
Pol I	Белок 43	FEN-1 или MF-1	Экзонуклеаза, удаляет РНК-затравки
РНКаза H	РНКаза H	РНКаза H1	Нуклеаза, удаляет РНК-затравки

В табл. 1.16 включены белки наиболее хорошо изученных систем репликации: *E. coli* и ее бактериофага T4, а также вируса SV40, размножающегося в культивируемых клетках человека (использованы общепринятые сокращения). При рассмотрении таблицы видно, что основные компоненты системы репликации ДНК в филогенезе функционально консервативны, и любой белковый компонент системы прокариот имеет свой прототип в системе репликации ДНК млекопитающих. Принимая во внимание только этот факт, можно ожидать наличие значительного сходства в механизмах репликации ДНК прокариотических и эукариотических организмов. Более удивительным представляется то, что у белков разных организмов, выполняющих одинаковые функции, в большинстве случаев отсутствует гомология в аминокислотных последовательностях. В частности, не обнаружено сходства у белка SSB (single-stranded DNA binding protein) *E. coli*, белкового продукта гена 32 фага T4 и белка RPA (replication protein A) репликативной системы человека. То же самое характерно и для β -субъединицы ДНК-полимеразы III (Pol III) *E. coli* (β -белок), белка 45 фага T4 и белка PCNA (proliferating cell nuclear antigen) человека. Это указывает на возможность выполнения одних и тех же функций полипептидными цепями с разными аминокислотными последовательностями, а также на вероятное конвергентное эволюционное происхождение таких белков и их функций из разных неродственных белков-предшественников.

4.1.2. Репликативная вилка *E. coli* и бактериофага T4

Во время редупликации ДНК ее дочерние синтезирующиеся цепи расходятся из точки репликации, образуя Y-подобную структуру, называемую *репликативной вилкой*. Именно в окрестностях этой точки разветвления и локализован функционирующий репликативный комплекс. Современные представления о строении репликативной вилки *E. coli* схематически изображены на рис. 1.46,а. В соответствии с этой моделью, ДНК-хеликаза перемещается в репликативной вилке впереди ДНК-синтезирующего белкового комплекса и расплетает цепи родительской ДНК, причем SSB-белок связывается с образующимися одноцепочечными участками, облегчая процесс расплетения.

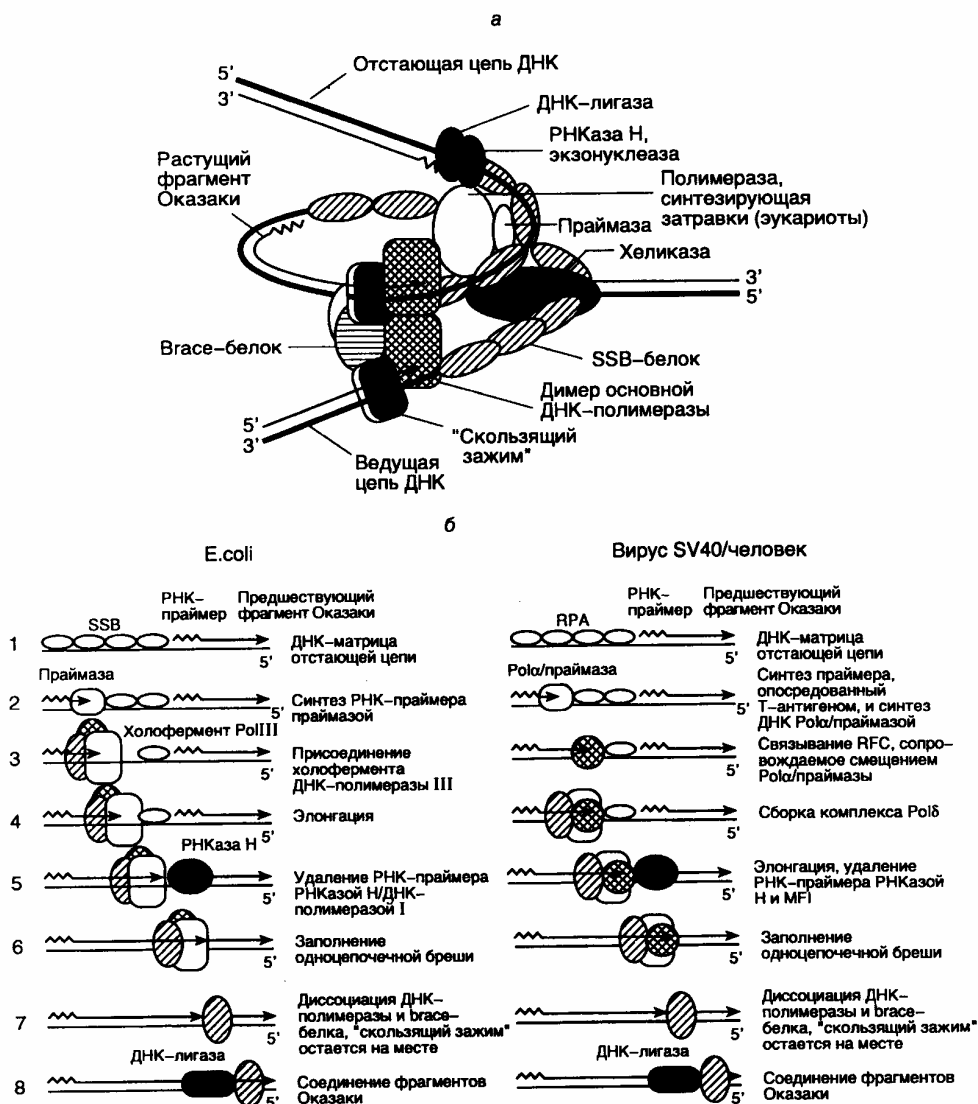


Рис. 1.46. Схема функционирующей репликативной вилки E. coli и эукариот и основных этапов репликации ДНК

а – репликативная вилка с основными белками репликации, в которой димер ДНК-полимеразы синхронно реплицирует обе цепи ДНК; **б** – этапы репликации (1–8) отстающей цепи ДНК. Этап 7 у эукариот является гипотетическим

Из-за антипараллельности и комплементарности цепей ДНК механизм репликации этих двух цепей существенно различается. Действительно, ДНК-полимераза обладает способностью синтезировать ДНК только в одном направлении: от 5'-конца к 3'-концу, перемещаясь вдоль ДНК-матрицы в направлении 3'→5'. В то же время комплементарные цепи ДНК противоположно

направлены (антипараллельны), и в силу своих свойств ДНК-полимераза не может реплицировать молекулу ДНК, просто перемещаясь от одного конца матричного дуплекса к другому. Для разрешения этого противоречия репликативный комплекс использует изящный механизм. На одной цепи ДНК синтез новой цепи происходит непрерывно, и образующаяся цепь называется *ведущей*, тогда как синтез другой цепи осуществляется прерывисто в виде коротких фрагментов, получивших название *фрагментов Оказаки* в честь ученого, впервые их открывшего. Эта вновь синтезируемая цепь ДНК называется *отстающей*. И хотя фрагменты Оказаки также синтезируются в направлении $5' \rightarrow 3'$, перемещение работающей ДНК-полимеразы вдоль матричной цепи ДНК при синтезе каждого индивидуального фрагмента Оказаки должно быть противоположным тому, которое имеет место в случае синтеза ведущей цепи. Образующиеся фрагменты Оказаки отстающей цепи далее соединяются друг с другом с помощью ДНК-лигазы.

Репликативный комплекс, который осуществляет синтез ведущей цепи ДНК, включает в себя минимальный (кор-) фермент ДНК-полимеразы III (белок 43 в случае фага Т4), подвижный связывающий β -белок с молекулярной массой 41 кДа ("sliding clamp", белок 45 у фага Т4) и белки γ -комплекса, состоящего из пяти полипептидов $\gamma\delta\delta'\chi\psi$ (скрепляющие белки – brace proteins). Функциональным аналогом белков γ -комплекса у бактериофага Т4 служит комплекс белковых продуктов генов 44/62. При облучении клеток *E. coli* УФ-светом в них индуцируется синтез укороченного β^* -белка (26 кДа), который является продуктом того же гена, что и β -субъединица холофермента ДНК-полимеразы III. По-видимому, функциональная роль β^* -белка заключается в обеспечении репликации ДНК на матрице, поврежденной УФ-светом.

При синтезе ведущей цепи ДНК репликативный комплекс *E. coli* функционирует весьма эффективно с высокой процессивностью. Напомним, что мерой процессивности является длина фрагмента вновь синтезированной макромолекулы, которую комплекс (или индивидуальные ферменты) способен образовывать в одном цикле, не диссоциируя от матрицы. Установлено, что холофермент ДНК-полимеразы III, в состав которого входят минимальный фермент (субъединицы α , θ и ϵ), β -белок, белки γ -комплекса и τ -белок, синтезирует ведущую цепь ДНК длиной в 50000 нуклеотидов со скоростью >500

нуклеотидов в секунду в одном цикле, ни разу не диссоциируя от ДНК-матрицы. Точность репликации ДНК холоферментом ДНК-полимеразы III поражает воображение. Частота ошибочных включений нуклеотидов не превышает 10^{-9} – 10^{-10} на нуклеотид за один раунд репликации. В то же время очищенные каталитические субъединицы реплицируют ДНК с пониженной точностью. В частности, изолированная α -субъединица допускает ошибки в опытах *in vitro* с частотой $\sim 6 \cdot 10^{-1}$ на нуклеотид за один раунд репликации. Частота ошибок, возникающих в ДНК, облученной УФ-светом, одна и та же в ведущей и отстающей цепях вновь синтезируемой ДНК *in vivo*.

Роль γ -комплекса заключается в распознавании РНК-затравок (праймеров) на матричной ДНК. γ -Комплекс связывается с единственным праймером ведущей цепи ДНК или с каждым из праймеров фрагментов Оказаки отстающей цепи, что, в свою очередь, делает возможным присоединение к промаркированным таким образом праймерам минимального фермента ДНК-полимеразы и β -белка (см. рис. 1.46,б).

Две молекулы β -белка входят в состав репликативного комплекса вслед за белками γ -комплекса, связываясь с ДНК позади белков γ -комплекса и оставляя 3'-конец праймера доступным для ДНК-полимеразы. Димер β -белка образует кольцо вокруг молекулы ДНК и стимулирует АТФазную активность белков γ -комплекса. Как уже упоминалось выше, функциональный аналог β -белка – продукт гена 45 бактериофага Т4, образует такую же пространственную структуру, охватывающую молекулу ДНК тремя молекулами. Молекулярная масса белка 45 составляет 2/3 от таковой β -мономера, и их аминокислотные последовательности негомологичны друг другу. Тем не менее, четвертичные структуры этих полипептидов и механизмы их функционирования обладают большим сходством.

β -Белки и белки γ -комплекса, будучи связанными с дуплексом праймер–матрица, обеспечивают присоединение к этому комплексу минимального фермента ДНК-полимеразы. Затем ДНК-полимераза при наличии доступных четырех дезоксирибонуклеозидтрифосфатов, используя праймер для инициации синтеза ДНК, с высокой эффективностью синтезирует цепь ДНК, комплементарную ДНК-матрице. Те же самые белки участвуют и в синтезе отстающей цепи ДНК. В этом случае прерывистый синтез ДНК многократно

инициируется на большом количестве праймеров, и ДНК синтезируется в виде фрагментов Оказаки длиной ~1000 нт. Синтез затравок, представляющих собой короткие последовательности РНК, обеспечивает продукт гена *dnaG* (белок 61 фага Т4). ДНК-полимераза начинает элонгацию цепей ДНК, присоединяя первый дезоксирибонуклеозидмонофосфат к 3'-концевому нуклеотиду РНК-затравки. В процессе элонгации участвуют β -белок и белки γ -комплекса, которые перемещаются вдоль молекулы ДНК вместе с каталитической субъединицей ДНК-полимеразы.

Ведущая и отстающая цепи ДНК реплицируются координированно, что обеспечивается димеризацией ДНК-полимеразных комплексов. В таком димере, который образуется при участии τ -белка, один ДНК-полимеразный комплекс осуществляет непрерывный синтез ведущей цепи ДНК, а другой – фрагментов Оказаки отстающей цепи. Для димеризации ДНК-полимеразы III *E. coli* необходим τ -белок, в то время как продукт гена 43 бактериофага Т4, по-видимому, изначально находится в виде димера. Другое различие репликативных комплексов *E. coli* и фага Т4 заключается в том, что холофермент ДНК-полимеразы *E. coli* (субъединицы $\alpha\epsilon\theta \cdot \gamma\delta\delta' \cdot \chi\psi \cdot \tau \cdot \beta$) сохраняется в виде стабильного комплекса и в отсутствие ДНК, тогда как холофермент Т4-ДНК-полимеразы существует только в присутствии матрицы.

Необычность ситуации во время синтеза ДНК в репликативной вилке заключается в том, что один и тот же белковый комплекс осуществляет как высоко процессивный непрерывный синтез ведущей цепи ДНК, так и прерывистый синтез фрагментов Оказаки отстающей цепи, претерпевая во втором случае периодическую диссоциацию от матрицы для инициации синтеза ДНК с каждого нового праймера. Для объяснения такого парадокса предположили, что холофермент ДНК-полимеразы способен узнавать 5'-конец каждого РНК-праймера, встречающегося после завершения синтеза очередного фрагмента Оказаки во время образования отстающей цепи ДНК в процессе репликации. Недавно с помощью оригинального экспериментального подхода удалось решить этот вопрос. В участок полипептидной цепи β -белка, контактирующий с минимальным ферментом ДНК-полимеразы III, методами генной инженерии ввели аминокислотную последовательность, узнаваемую и фосфорилируемую протеинкиназой. Измеряя скорость фосфорилирования этих

сайтов в условиях избытка протеинкиназы во время синтеза ДНК *in vitro*, определили кинетику ассоциации и диссоциации комплексов β -белок–ДНК-полимераза по изменению уровня защищенности сайтов фосфорилирования от действия протеинкиназы. Полученные результаты интерпретировали таким образом, что во время связанного с синтезом фрагментов Оказаки перемещения минимального фермента ДНК-полимеразы и γ -комплекса вдоль одноцепочечной ДНК-матрицы, покрытой SSB-белком, оба белка прочно связаны с β -белком и матрицей. При встрече репликативного белкового комплекса с дуплексом, образованным матричной ДНК и РНК-затравкой, β -белок остается связанным с вновь синтезированной ДНК, а у отделившихся ДНК-полимеразы и белков γ -комплекса появляется возможность вступить в новый цикл синтеза фрагмента Оказаки путем взаимодействия с очередным дуплексом РНК-затравка–матрица. При этом вхождение ДНК-полимеразы в новый репликативный комплекс облегчается наличием в нем β -белка и белков γ -комплекса, ассоциированных с очередным РНК-праймером. Таким образом, холофермент ДНК-полимеразы III обладает способностью распознавать молекулярное окружение, создаваемое матричной ДНК, осуществлять терминацию синтеза ДНК при наличии сигнала в виде дуплекса ДНК–затравка и реинициировать синтез ДНК на следующем праймере. В итоге одна и та же молекула ДНК-полимеразы III в составе реплицирующего комплекса способна проводить синтез всех фрагментов Оказаки отстающей цепи реплицируемой ДНК, последовательно осуществляя инициацию, терминацию и реинициацию синтеза каждого из них.

После очередной терминации синтеза ДНК отстающей цепи 3'-конец вновь синтезированной ДНК оказывается вплотную приближенным к 5'-концу праймера следующего фрагмента Оказаки. Для соединения двух фрагментов с помощью ДНК-лигазы необходимы предварительное удаление РНК-праймера и достройка цепи ДНК в образующейся брешу. РНК-затравка удаляется с помощью РНКазы H, нуклеазы, специфически расщепляющей РНК в ДНК–РНК-гибридах, и(или) с участием 5'→3'-экзонуклеазы ДНК-полимеразы I. Во втором случае одновременно с удалением праймера происходит застройка образующейся брешу той же ДНК-полимеразой. В итоге два соседних фрагмента Оказаки вплотную приближаются друг к другу и оказываются

отделенными лишь одноцепочечным разрывом, который может репарироваться ДНК-лигазой. В настоящее время остается открытым вопрос о механизмах координации удаления РНК-праймеров из фрагментов Оказаки с самим процессом репликации ДНК.

Кроме вышеупомянутых дуплексов праймеры–ДНК, холоферменты бактериальных и фаговых ДНК-полимераз, по-видимому, способны адекватно реагировать на другие стерические препятствия, возникающие на пути следования вдоль реплицируемой молекулы ДНК. В частности, ДНК-полимераза фага Т4 в процессе репликации может расходиться с транскрибирующими ту же ДНК молекулами РНК-полимеразы, не диссоциируя из репликативного комплекса и не вытесняя РНК-полимеразу с матрицы. Кроме того, репликативный комплекс может распознавать повреждения ДНК, возможно, маркированные специфическими белками, и прекращать репликацию соответствующего участка, останавливаясь или диссоциируя от матрицы. Репликация таких участков ДНК возобновляется после ликвидации повреждений ферментами репаративной системы. Для диссоциировавшей ДНК-полимеразы это становится возможным благодаря тому, что 3'-конец вновь синтезированной цепи ДНК в месте остановки репликации остается связанным с β -белком, который облегчает повторное вхождение диссоциировавшей ДНК-полимеразы в репликативный комплекс.

4.1.3. Особенности функционирования репликативной вилки эукариот

Механизмы репликации ДНК у высших эукариот менее изучены из-за их большей сложности. Основные результаты получены на модельной системе с ДНК вируса SV40, в которой процесс репликации исследовали в зараженных клетках человека, культивируемых *in vitro*. В этой системе вирусный белок, называемый Т-антигеном, выполняет многие функции, необходимые для репликации вирусной ДНК. Во-первых, он является белком-инициатором, необходимым для инициации репликации; во-вторых, он обладает ДНК-хеликазной активностью, т.е. расплетает цепи реплицируемой ДНК перед работающей ДНК-полимеразой, и, в-третьих, Т-антиген необходим для правильного взаимодействия с ДНК ферментного комплекса, синтезирующего праймеры (*праймсомы*). Тем не менее, вирус SV40 использует для репликации

ДНК своей небольшой хромосомы и многие белки клетки-хозяина, что позволяет исследовать функционирование репликативного комплекса клеток человека в такой относительно простой системе.

У эукариот обнаружены шесть ДНК-полимераз, три из которых – α , δ и ε – непосредственно участвуют в репликации хромосомной ДНК (табл. I.17). Аминокислотные последовательности этих трех ферментов гомологичны друг другу и последовательности продукта гена 43 бактериофага T4. Эукариотическая ДНК-праймаза в отличие от аналогичного белка прокариот образует постоянный комплекс с ДНК-полимеразой α , роль которого, по-видимому, ограничивается синтезом праймеров при репликации обеих цепей ДНК.

Белок PCNA и фактор репликации C (RFC) также образуют стабильный комплекс с ДНК-полимеразой δ , а в определенных условиях стимулируют и активность ДНК-полимеразы ε . Во многих отношениях PCNA и RFC являются функциональными аналогами соответственно β -белка и белков γ -комплекса *E. coli* (см. рис. I.46,б), и их роль в синтезе ведущей и отстающей цепей ДНК вируса SV40 хорошо известна. Механизмы репликации ДНК прокариот и эукариот существенно различаются в том отношении, что во втором случае синтез ведущей и отстающей цепей ДНК

Таблица I.16

Эукариотические ДНК-полимеразы и их функциональные гомологи у прокариот

ДНК-полимеразы	Ген дрожжей	Гомолог E. coli	Молекулярные массы субъединиц, кДа	Биологические функции
α	<i>POL1</i>	?	160–185	Синтез ведущей цепи геномной ДНК в репликативной вилке; в комплексе с праймазой обеспечение синтеза праймеров на обеих цепях ДНК
β		Pol I	40	Заполнение брешей при эксцизионной репарации ДНК, участие в рекомбинации
γ	<i>MIP1</i>	–	140 (человек) 116 (дрожжи)	Репликация митохондриальной ДНК
δ	<i>POL3</i>	Pol III	125	Синтез отстающей цепи геномной ДНК в репликативной вилке
ε	<i>POL2</i>	Pol II (?)	210–230	Репарация ДНК, регуляция клеточного цикла (?)
ζ	<i>REV3</i> и <i>REV7</i>	Pol IV (DinB/P)	173 и 29	Синтез ДНК на поврежденной матрице при SOS-ответе
η	<i>RAD30</i>	DinB, UmuC	70	Синтез ДНК на поврежденной матрице, с включением остатков А напротив тиминовых димеров

Примечание. ? – гомологи неизвестны.

осуществляют разные ДНК-полимеразы (α и δ соответственно), тогда как у *E. coli* обе цепи ДНК синтезируются димером ДНК-полимеразы III. ДНК-полимераза α проводит инициацию синтеза ведущей цепи в точках начала репликации, а ДНК-полимераза δ осуществляет циклические реинициации синтеза фрагментов Оказаки, по-видимому, распознавая наличие 5'-концевого нуклеотида очередного праймера с последующей диссоциацией от матричной ДНК и присоединением к ней для реинициации синтеза следующего фрагмента Оказаки. Созревание фрагментов Оказаки у эукариот требует удаления РНК-затравок с помощью 5'→3'-экзонуклеазы (белковые факторы FEN-1 или MF-1) и РНКазы H1, а также ковалентного соединения фрагментов друг с другом под действием ДНК-лигазы I.

Роль ДНК-полимеразы ε в настоящее время не ясна. Возможно, этот фермент непосредственно участвует в репликации или в сопряженной с репликацией репарации повреждений ДНК, а также в регуляции клеточного цикла.

ДНК-полимераза ζ обнаружена в 1996 г. у дрожжей *S. cerevisiae*. При исследовании белков Rev3 и Rev7, которые необходимы для мутагенеза, индуцируемого в ответ на повреждения ДНК, оказалось, что их комплекс обладает ДНК-полимеразной активностью. Эта полимераза способна эффективно использовать в качестве матрицы ДНК, содержащую циклобутановые димеры. В таких условиях активность ДНК-полимеразы α составляет лишь 1% от активности ДНК-полимеразы ζ .

ДНК-полимераза η , так же как и предыдущий фермент, участвует в SOS-ответе дрожжей на генотоксические воздействия. В присутствии всех четырех дезоксирибонуклеозидтрифосфатов она осуществляет включение в строящуюся цепь ДНК напротив тиминовых димеров только правильных нуклеотидов (А). Подробнее о функциях бактериальных гомологов двух последних ДНК-полимераз в SOS-мутагенезе см. в разделе 5.1.2.

4.2. Регуляция репликации ДНК

Подробное рассмотрение молекулярных механизмов регуляции репликации ДНК выходит за рамки книги, поэтому ограничимся несколькими замечаниями по данному вопросу и более детально обсудим лишь механизм

регуляции репликации у *E. coli*, в том числе и бактериальных плазмид, что имеет непосредственное отношение к функционированию плазмидных векторов в бактериальных клетках.

Синтез ДНК тесно связан с другими процессами, подготавливающими деление клеток, так как передача необходимой генетической информации родительских клеток дочерним является для клеток-потомков жизненно важной. Наличие избыточной генетической информации отрицательно сказывается на жизнеспособности клеток, тогда как недостаток ее, возникающий вследствие недорепликации ДНК, приводит к летальному эффекту из-за отсутствия жизненно важных генов. Однако процесс передачи генетической информации от родительских клеток дочерним у эукариот не ограничивается простой редупликацией ДНК хромосом. Так, для насекомых многих видов характерно наличие гигантских *политенных* хромосом, которые возникают в результате множественных раундов репликации ДНК исходных хроматид, не сопровождающейся их расхождением.

Политенизация хромосом представляет обширный класс генетических явлений, связанных с избирательной избыточной репликацией (*мультипликацией*) или недорепликацией отдельных генетических локусов эукариот. Ярким примером такого рода является изменение числа генов рибосомных РНК у животных. Амплификация генов рРНК в ооцитах амфибий происходит путем образования их внехромосомных (экстрахромосомных) копий в виде кольцевых молекул рибосомных (р) ДНК, которые далее реплицируются по механизму "катыщегося кольца". При этом в каждой клетке амплифицируется только по одному из сотен повторов рДНК, так что амплификация рДНК на одном повторе каким-то образом подавляет процесс амплификации на других, и все образовавшиеся повторы одного ооцита идентичны, но отличаются от наборов амплифицированных рДНК других ооцитов. Строгая стадие- и тканеспецифичность, а также избирательная амплификация только одного повтора рДНК указывают на наличие тонких регуляторных механизмов процесса репликации и в этом случае.

Характерными примерами возрастания числа генов вследствие их избирательной репликации являются *магнификация* генов рРНК и изменение числа генов, определяющих устойчивость клеток к лекарственным препаратам. В первом случае утрата части генов рРНК у дрозофилы в результате делеции

сопровождается постепенным восстановлением их числа, тогда как во втором случае у клеток, находящихся в условиях селективного действия токсичного для них лекарственного препарата, возрастает число копий генов, необходимых для его нейтрализации. В частности, это характерно для гена дигидрофолатредуктазы в присутствии метотрексата. Высказывается предположение, что в основе изменения числа копий таких генов лежит механизм неравного кроссинговера.

Репликация хромосом бактерий тесно сопряжена с метаболизмом клеток. Например, частота инициаций новых раундов репликации зависит от скорости роста бактериальных клеток, и в клетках быстро растущих бактерий могут содержаться хромосомы с несколькими работающими репликативными вилками, хотя для репликации одной бактериальной хромосомы их требуется только две, инициированные в единственной области начала репликации (*ori*) и расходящиеся в противоположных направлениях. Это позволяет бактериям при благоприятных условиях затратить для генерации меньше времени, чем для полной репликации бактериальной хромосомы. Очевидно, что для поддержания строго упорядоченного характера репликации должны существовать тонкие механизмы регуляции репликации на уровне инициации новых раундов. Такие механизмы, действительно, существуют.

Наиболее хорошо изученными в настоящее время являются механизмы регуляции синтеза ДНК у *E. coli*, в том числе механизмы контроля числа копий у небольшой плазмиды *E. coli* ColE1, которые будут рассмотрены ниже более подробно из-за важности этих явлений для генной инженерии.

4.2.1. Инициация репликации ДНК у *E. coli* и ее регуляция

Репликация хромосомной ДНК у бактерий играет ключевую роль в их жизненном цикле. В ходе этого процесса микроорганизмы редулицируют свой геном, а образовавшиеся дочерние геномы далее переходят в дочерние клетки. Высокая точность, с которой бактерии осуществляют такие процессы, указывает на наличие специальных механизмов их координации и контроля.

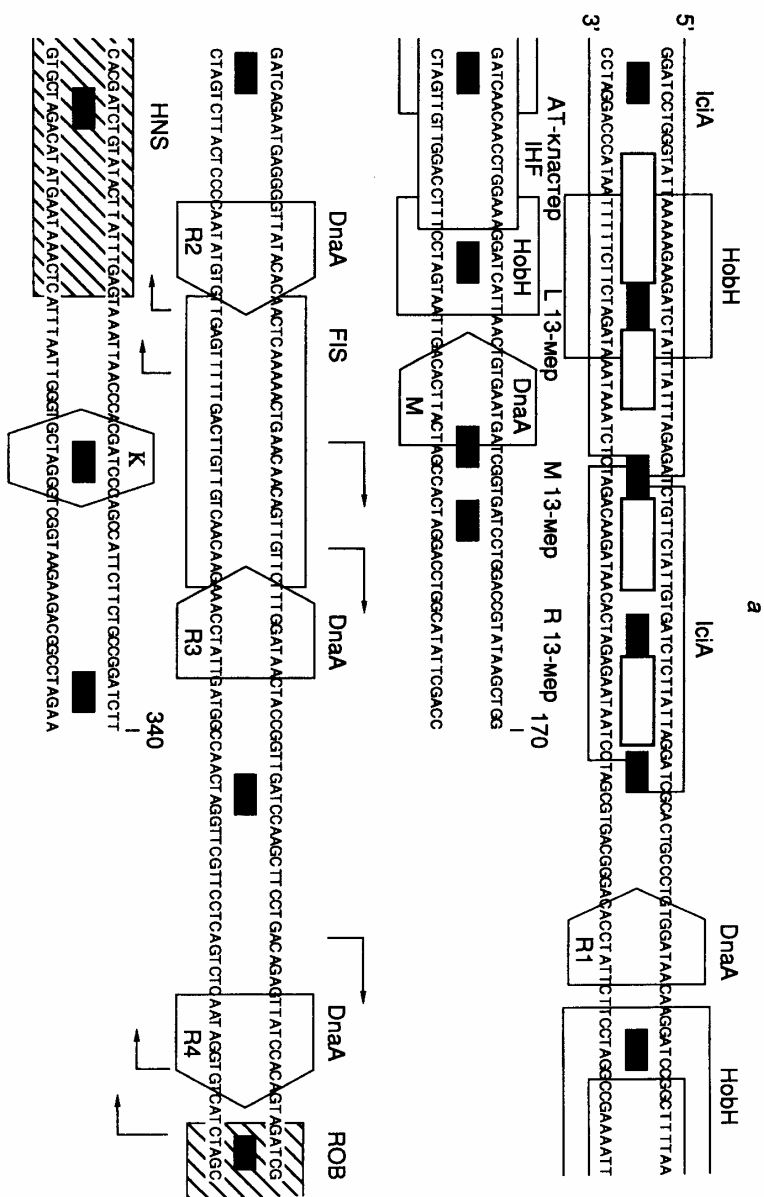
Структура области начала репликации *oriC*. Хромосома *E. coli* содержит единственную *область начала репликации* (*origin*), названную *oriC*, на которой происходит инициация репликации (рис. 1.47,а). Размер минимальной области начала репликации, обеспечивающей автономную

репликацию хромосомы, составляет 258 п.о. (положение 11–268 на рис. 1.47). Сравнение первичных структур областей начала репликации различных энтеробактерий показало, что их последовательности представлены короткими консервативными участками, которые перемежаются дивергировавшими сегментами ДНК, длины которых, однако, высококонсервативны. Консервативные участки оказались сайтами связывания регуляторных белков, разделенных спейсерными последовательностями. *OriC* содержит пять консенсусных 9-нуклеотидных сайтов связывания инициатора DnaA (непалиндромные повторы), названных DnaA-боксами. У всех энтеробактерий области начала репликации содержат 9–14 сайтов GATC, положение восьми из которых консервативно.

В левой части *oriC* находится АТ-богатая область, содержащая три похожих последовательности длиной в 13 нуклеотидов, каждая из которых начинается с GATC. Здесь же локализован АТ-кластер, который вместе с левой 13-нуклеотидной последовательностью образует область нестабильной спирали ДНК (*ДНК-расплетающий элемент*). Этот участок ДНК может быть заменен без потери функции на аналогичный по нуклеотидному составу, но с другой последовательностью нуклеотидов.

OriC содержит сайты связывания белков, изгибающих ДНК, IHF (integration host factor) и FIS (factor for inversion stimulation). Оба белка, по-видимому, помогают инициатору DnaA раскручивать ДНК.

Димерный белок IciA, состоящий из субъединиц с молекулярной массой 33 кДа, специфически связывается с АТ-богатыми 13-мерными повторами. Функция этого белка неизвестна, так же как и функция белка Rob, который специфически взаимодействует с 26-нуклеотидным сайтом в правой части DnaA-бокса R4. ДНК вблизи Rob-сайта обнаруживает изгиб, который более ярко выражен у молекул, полностью метилированных Dam-метилтрансферазой (см. ниже). С такими полностью метилированными ДНК взаимодействует гистоноподобный белок H-NS, сайт связывания которого перекрывается с Rob-сайтом. Это взаимодействие оказывает влияние на функционирование *oriC*.



б



Рис. 1.47. Структура области начала репликации хромосомы *E. coli* (а) и схема инициации ее репликации (б)

HobN – белок, взаимодействующий с метилированной по одной цепи ДНК области начала репликации (hemimethylated origin binding)

Функции белка DnaA. Белок DnaA играет ключевую роль в сборке *реплисомы* – многокомпонентного белкового комплекса, осуществляющего двунаправленный синтез ДНК. Белок распознает область начала репликации и привлекает к месту сборки остальные белковые компоненты реплисомы.

Этапы инициации синтеза ДНК на *oriC*. Сборка *исходного комплекса* начинается с взаимодействия белка DnaA с DnaA-боксами R1–R4 и M (см. рис. 1.47,б). Для успешного прохождения последующих этапов сборки реплисомы белок DnaA должен находиться в комплексе с ATP и взаимодействовать с сверхспирализованным *oriC*. С помощью электронного микроскопа исходный комплекс обнаруживается в виде компактной эллипсоидной структуры, содержащей ~20 мономеров DnaA, которая закрывает *oriC*. Исходный комплекс обладает высокоупорядоченной структурой.

В присутствии ATP в высокой концентрации (5 mM) исходный комплекс превращается в *открытый комплекс*. В этом комплексе происходит частичное расплетение AT-богатых 13-нуклеотидных повторов, расположенных в левой

части *oriC*. При 37° или выше единственный белок DnaA может обеспечивать расплетение ДНК. Для образования открытого комплекса при более низких температурах требуется участие структурирующего белка HU или интеграционного фактора бактерии-хозяина IHF. В открытом комплексе обнаруживают небольшие участки расплетенной ДНК в правой части *oriC* между DnaA-боксами R2 и R4, которые рассматривают как места посадки хеликазы.

Белок DnaB является хеликазой репликативной вилки и входит в открытый комплекс с образованием *препраймирующего комплекса I*, взаимодействуя с одноцепочечными участками частично расплетенной ДНК. Такие участки подготавливаются белком DnaA, который вытесняет SSB-белок с соответствующих сайтов. DnaB входит в препраймирующий комплекс I в виде гексамеров, образовавших комплекс с шестью мономерами DnaC, каждый из которых связывает одну молекулу АТФ. В этом комплексе хеликазная активность белка DnaB блокирована. Освобождение DnaC из комплекса происходит в результате гидролиза АТФ. Следствием этого является активация хеликазы DnaB и ее правильное расположение в комплексе. Совокупность этих событий превращает препраймирующий комплекс I в *препраймирующий комплекс II*.

Хеликаза должна начать функционировать в месте старта репликативной вилки в правой части *oriC* вблизи DnaA-боксов R2, R3 и R4. Для этого она должна быть транслоцирована от места ее первоначального вхождения в комплекс к точке начала репликации. Предполагается, что транслокация ассоциирована с АТФ-зависимым освобождением из комплекса белка DnaC, что сопровождается активацией хеликазы.

В *праймирующем комплексе* хеликаза DnaB взаимодействует с DnaG-праймазой, которая играет ключевую роль в обеспечении инициации репликации именно на *oriC*. Оба этих фермента обеспечивают сопряжение функционирования двух репликативных вилок, движущихся в противоположные стороны. В бесклеточной системе при низких концентрациях праймазы репликация становится однонаправленной и может иницироваться не на *oriC*. В праймирующем комплексе присутствие белка DnaA больше не требуется, и он после освобождения из комплекса может быть повторно использован для инициации репликации на другом *oriC*. Полагают, что во время

координированной сборки двух репликативных вилок в одной из них синтезируется праймер, который становится затравкой при синтезе ведущей цепи другой репликативной вилкой, движущейся в противоположном направлении. Праймаза в праймирующем комплексе функционирует по дистрибутивному механизму. После синтеза праймеров она покидает репликативную вилку и заменяется новой молекулой праймазы во время образования очередного фрагмента Оказаки.

При образовании реплисомы в каждой репликативной вилке происходит АТР-зависимое формирование димерного комплекса холофермента ДНК-полимеразы III, связанного с 3'-концами праймеров (скользящий зажим, см. выше). Вслед за этим происходит координированная элонгация праймеров, сопровождаемая двунаправленным синтезом ведущих и отстающих цепей ДНК. В бесклеточной системе точки начала синтеза ведущих цепей локализованы в *oriC* вблизи DnaA-боксов R2, R3 и R4.

Механизмы контроля инициации репликации *in vivo*. Инициация репликации ДНК у *E. coli* регулируется, по крайней мере, на трех уровнях: 1) инициация синхронизирована с клеточным циклом; 2) синтез ДНК в каждой области начала репликации в клеточном цикле иницируется только один раз; 3) инициация происходит синхронно во всех областях начала репликации, присутствующих в данной бактериальной клетке. Установлено, что синтез ДНК начинается после того, как масса бактериальной клетки в расчете на одну область начала репликации достигает определенного значения, названного *массой инициации* (initiation mass). В качестве основного водителя ритма (пейсмекера), играющего ключевую роль в контроле инициации репликации, в настоящее время рассматривается белок DnaA.

Подавление синтеза белка *in vivo* сопровождается завершением уже иницированного синтеза ДНК на фоне прекращения новых раундов инициации. Возобновление синтеза белка приводит к инициации репликации после лаг-периода в одну клеточную генерацию. При наличии всех необходимых белков инициация чувствительна к рифампину – специфическому ингибитору бактериальной РНК-полимеразы, что указывает на зависимость инициации от синтеза нетранслируемой РНК.

Роль топологии *oriC* в инициации репликации. Топоизомераза I и топоизомераза II (ДНК-гираза) поддерживают бактериальную хромосому в

негативно суперскрученном состоянии. Приблизительно половина супервитков нейтрализуется гистоноподобными белками HU, IHF и FIS, тогда как остающаяся сверхспирализация бактериальной хромосомы облегчает транскрипцию, репликацию и сайт-специфическую рекомбинацию. Предполагается, что бактериальная хромосома состоит из 40–50 суперскрученных доменов с ~25 супервитками на 1 т.п.о. ДНК. В настоящее время отсутствуют точные данные о топологическом состоянии *oriC*, необходимом для инициации репликации у *E. coli*. Известно, что мутации в гене топоизомеразы *topA* супрессируют температурно-чувствительные мутации *dnaA(Ts)*. Предполагается, что в этих мутантных штаммах топология *oriC* изменена таким образом, что допускает инициацию репликации при меньших внутриклеточных концентрациях белка DnaA. Кроме того, на важность определенного топологического состояния *oriC* для инициации указывает факт нарушения инициации у мутантных бактерий с измененным геном *gyrB(Ts)*, кодирующим В-субъединицу ДНК-гиразы.

Активация репликации транскрипцией. В том случае, если сверхспирализация минихромосом или плазмид, содержащих *oriC*, недостаточна для инициации их репликации, инициация может происходить при одновременной транскрипции ДНК в окрестностях *oriC*. Изменение топологии *oriC* в этом случае может осуществляться за счет образования *R-петель* (ДНК–РНК-гибрида в двухцепочечной ДНК) или вследствие транскрипции, как таковой, при которой перед транскрибирующей РНК-полимеразой имеет место локальная положительная сверхспирализация ДНК, а вслед за ней – отрицательная. Это облегчает образование открытых комплексов при инициации синтеза ДНК.

Роль белка DnaA в регуляции инициации репликации. ~60 минут необходимо бактерии для репликации хромосомной ДНК, разделения дочерних хромосом и подготовки к новому делению. Следовательно, клетки со временем генерации короче этого периода (например при повышенных температурах на богатых питательных средах) должны иницировать репликацию хромосом, предназначенных для последующих делений, до завершения предыдущего раунда репликации. Таким образом, в отдельной клетке может содержаться реплицирующаяся хромосома со множественными точками начала репликации. При этом инициация репликации на множественных областях начала

репликации происходит одновременно.

Сверхпродукция DnaA в бактериях приводит к резкому возрастанию частоты инициаций репликации без изменения общей скорости синтеза ДНК, что указывает на DnaA как на позитивный регулятор этого процесса. Среди моделей, объясняющих механизм регуляторного действия белка DnaA наибольшее распространение получила модель титрования DnaA. В соответствии с этой моделью весь вновь синтезируемый белок DnaA связывается (титруется) DnaA-боксами *oriC* хромосомы. Как только количество молекул инициатора превышает число внутриклеточных DnaA-боксов (все DnaA-боксы оказываются занятыми белком), происходит инициация синтеза ДНК. После запуска инициации на одном *oriC* наблюдается освобождение молекул DnaA, резкое повышение его внутриклеточной концентрации и синхронная инициация синтеза ДНК на других доступных областях начала репликации. При этом ассоциация с мембранами первой *oriC* защищает ее от использования в реинициации.

Роль Dam-метилирования в инициации синтеза ДНК. Как уже упоминалось выше, Dam-метилтрансфераза *E. coli* модифицирует остатки аденина в последовательностях 5'-GATC. В результате репликации молекула ДНК временно превращается из полностью метилированной молекулы в метилированную по одной цепи, что позволяет клетке распознавать вновь синтезированную ДНК. Расположение кластеров Dam-сайтов в *oriC* энтеробактерий высококонсервативно (см. рис. 1.47,а). Неметилированная или наполовину метилированная плазмидная ДНК в клетках *dam*-мутантов не реплицируется, хотя и служит субстратом в бесклеточной системе репликации. Репликация хромосомной ДНК у *dam*-мутантов начинается на *oriC*, однако контроль репликации нарушен, что проявляется в асинхронности репликации на множественных *oriC*. Оказалось, что лишь наполовину метилированная, но не полностью метилированная или неметилированная *oriC*-ДНК специфически связывается с фракцией мембран *E. coli* *in vitro*. При этом в быстро растущих клетках ~1/3 времени генерации *oriC*-ДНК находится в наполовину метилированном состоянии, после чего полностью метилируется. То же самое характерно и для промотора гена инициатора DnaA, у которого метилированное наполовину состояние связано с подавлением транскрипции гена. В отличие от этого реметилирование вновь синтезированной цепи ДНК остальной части

бактериальной хромосомы происходит быстро – в течение 1–2 мин. На основании такого рода данных высказывается предположение, что в не полностью метилированном состоянии вышеупомянутые последовательности экранированы бактериальными мембранами от контактов с регуляторными белками и не могут участвовать в повторном раунде инициации репликации (период *эклипса*). Мутации в гене *seqA* резко уменьшают время *эклипса*, что проявляется в асинхронности инициаций репликации. Белок SeqA оказался негативным регулятором инициации репликации, действующим на этапе взаимодействия *oriC* с бактериальными мембранами.

Роль белка SeqA в регуляции репликации бактериальных хромосом.

Ген *seqA* кодирует белок длиной в 181 аминокислотный остаток, инактивация которого летальна для бактериальных клеток. Исследование взаимодействия этого белка с неметилированной, частично и полностью метилированной областями начала репликации методом смещения полос при электрофорезе в полиакриламидном геле показало его предпочтительное связывание с частично метилированными последовательностями. Однако для полной (контекст-зависимой) специфичности его взаимодействия требуется присутствие дополнительных факторов. Действительно, в составе ДНК-белковых комплексов, образованных с участием частично метилированных последовательностей *oriC*, обнаружен белок с молекулярной массой 24 кДа, который специфически взаимодействует с метилированной цепью ДНК в *oriC*. Скрининг клонотеки последовательностей *E. coli* позволил клонировать ген *hobH* (hemimethylated origin binding), кодирующий этот белок. Мутации по данному гену приводили к частичной утрате бактериальными клетками синхронизации в инициациях репликации, что также косвенно указывает на участие белка HobH в регуляции инициации репликации бактериальных хромосом на ранних стадиях клеточного цикла. Однако истинная роль этого белка в репликации окончательно не известна.

Период *эклипса* может заканчиваться в результате постепенного завершения метилирования частично метилированной последовательности *oriC*, находящейся в комплексе с мембранами. Полное метилирование этих последовательностей предотвращает их взаимодействие с мембранами и делает доступными для инициатора DnaA.

Терминация репликации. Встреча двух репликативных вилок в конце

цикла репликации бактериальной хромосомы сопровождается несколькими событиями, которые необходимы для полного разделения двух образовавшихся бактериальных хромосом до деления клетки. Движение репликативных вилок навстречу друг другу сопровождается гомологичной рекомбинацией между дочерними хроматидами. В том случае, если количество произошедших рекомбинаций нечетное, образуется димер бактериальной хромосомы, тогда как при четном числе рекомбинаций – две катенированные (зацепленные друг за друга) хромосомы. Во втором случае разделение катенов с помощью топоизомеразы IV приводит к полному разделению дочерних хромосом, тогда как в случае димера бактериальной хромосомы этого недостаточно. Разделение димера с образованием мономеров происходит в результате сайт-специфической рекомбинации в локусе *dif* под действием резольвазы (сайт-специфической рекомбиназы) XerCD.

4.2.2. Регуляция репликации плазмиды ColE1

Многие клетки прокариот в дополнение к основной хромосоме содержат небольшие внехромосомные ДНК, называемые *плазмидами*. Плазмиды, размеры которых варьируют от нескольких тысяч до сотен тысяч пар оснований, а число копий на клетку – от одной до нескольких сотен, способны к автономной (независимой от основной хромосомы) репликации и стабильно наследуются в ряду клеточных поколений. Хотя многие плазмиды дают клеткам-хозяевам ощутимые селективные преимущества (устойчивость к антибиотикам, тяжелым металлам и т.п.), большинство из них являются *криптическими*, т.е. не проявляющимися в видимом фенотипе клеток. Поскольку их существование – это весома нагрузка на метаболизм клеток-хозяев, остается непонятным смысл их эволюционной стабильности. Несмотря на то что в природных условиях бактериальные клетки, по-видимому, не испытывают давления отбора, направленного на сохранение плазмид внутри клеток, последние с помощью тонких механизмов, регулирующих число их копий в клетках, стабильно сегрегируют между дочерними бактериальными клетками.

Область начала репликации небольшой плазмиды ColE1, несущей гены устойчивости к колицинам, традиционно используется в генной инженерии при конструировании векторных молекул ДНК, которые находят применение для

клонирования и экспрессии в клетках *E. coli* коротких последовательностей нуклеотидов. Именно поэтому целесообразно рассмотреть механизмы контроля репликации плазмиды ColE1.

Инициация репликации плазмиды ColE1. Репликация плазмиды ColE1 происходит в одном направлении (однонаправленная репликация) с использованием репликативного аппарата клетки-хозяина. Сама по себе плазида не кодирует ни одного фермента, который требовался бы для ее репликации. Область начала репликации содержит два промотора, один из них обеспечивает синтез РНК-праймера (РНК II), необходимого для инициации репликации плазмиды. Синтезированная РНК II, длина которой зависит от типа реплицируемой плазмиды, далее подвергается процессингу с помощью РНКазы H с образованием РНК длиной в 550 нуклеотидов. Эта молекула эффективно используется ДНК-полимеразой I в качестве праймера при синтезе ведущей цепи ДНК. В отсутствие РНКазы H затравкой во время репликации служит 3'-конец РНК II, хотя и с меньшей эффективностью. В клетках, дефектных по РНКазе H и ДНК-полимеразе, инициация репликации ColE1 осуществляется ДНК-полимеразой III с участием РНК II по механизму, подробно рассмотренному выше.

Все три механизма инициации репликации плазмиды основаны на уникальном свойстве РНК II образовывать стабильный ДНК–РНК-гибрид в области начала репликации. Действительно, обычные транскрипты освобождаются из транскрипционного комплекса после завершения транскрипции и отделения РНК-полимеразы от матрицы, чего не происходит с РНК II. Анализ мутантов плазмиды, дефектных по репликации, а также их ревертантов показал, что в стабильном гибриде РНК II с матрицей происходит взаимодействие между G-богатой петлей РНК II, образованной 265 нуклеотидами выше точки инициации репликации (положение –265), и C-богатым участком ДНК, расположенным в окрестностях нуклеотида –20 (рис. 1.48,а). Обе эти последовательности оказались консервативными у родственных плазмид рMB1, р15A и KSF1030. Взаимодействия между указанными последовательностями, по-видимому, происходят в тот момент, когда РНК-полимераза еще находится в транскрипционном комплексе и цепи ДНК в окрестностях комплекса расплетены. Равновесие между двумя альтернативными конформациями РНК II является критическим в определении

доли молекул РНК, остающейся в ДНК–РНК-гибриде, необходимом для инициации репликации плазмиды. Выбор между двумя альтернативными конформациями РНК II определяется первичной структурой участка, расположенного между нуклеотидами –359 и –380 (последовательность β) (см. рис. 1.48,б). Эта последовательность может взаимодействовать с выше расположенной комплементарной последовательностью α (структура $\alpha\beta$) или с гомологичной последовательностью γ , расположенной ниже (структура $\beta\gamma$). После того как РНК-полимераза транскрибирует первые 200 нуклеотидов, образовавшаяся РНК II формирует временную вторичную структуру, для которой характерно наличие трех доменов типа "стебель–петля" (I, II и III). Удлинение РНК II еще на несколько нуклеотидов приводит к разрушению стебля III и образованию стебля IV, который стабилизируется в результате комплементарных взаимодействий между последовательностями α и β . В течение последующей элонгации РНК II у нее возникают две альтернативные возможности формировать свою вторичную структуру. Выбор в пользу той или иной конформации зависит от того, останется ли последовательность β связанной с последовательностью α или же образует новые контакты с γ -последовательностью. Переход от комплементарных пар $\alpha\beta$ к $\beta\gamma$ сопровождается сильными изменениями конформации РНК II, которые в конечном счете определяют ее способность служить праймером при репликации плазмиды. Молекулы РНК II в конформации $\alpha\beta$ могут образовывать РНК–ДНК-гибрид, служащий субстратом для РНКазы H, а в конформации $\beta\gamma$ такой способностью не обладают. Предложенная модель подтверждается, прежде всего, тем, что мутации, делающие предпочтительным образование конформации $\beta\gamma$ из-за дестабилизации стебля IV, затрудняют функционирование РНК II в качестве праймера и приводят к понижению числа копий плазмиды ColE1 внутри бактериальных клеток. Такие мутантные плазмиды, дефектные по репликации, активизируются в результате супрессорных мутаций, стабилизирующих стебель IV. Таким образом, инициация репликации плазмиды ColE1 зависит от способности РНК II образовывать РНК–ДНК-гибрид вблизи точки начала репликации (*ori*). При этом на образование гибрида оказывают влияние вторичная и третичная структуры выше расположенной последовательности нуклеотидов предшественника

а – предполагаемая вторичная структура РНК II, после транскрибирования РНК-полимеразой ~ 500 нуклеотидов ДНК плазмиды; дальнейшее удлинение РНК II сопровождается образованием ДНК–РНК-гибрида (жирная стрелка) между РНК II и транскрибируемой ДНК;

б – возможный механизм контроля репликации плазмиды. В верхней части рисунка изображена генетическая карта участка ДНК, необходимого для инициации репликации плазмидной ДНК и ее контроля. Схематически представлены пространственные структуры двух ингибиторов репликации плазмиды: РНК I и белка Rop. В нижней части изображены две альтернативные конформации РНК II, образующиеся под действием РНК I, I–X – элементы вторичной структуры

Контроль числа копий плазмиды ColE1. Контроль инициации репликации плазмиды ColE1 осуществляется главным образом на уровне изменения пространственной структуры РНК II. Поскольку плазмиды контролируют собственный биосинтез, т.е. их репликация проходит по аутокаталитическому механизму, было постулировано, что инициация репликации ColE1 находится под влиянием ингибитора, кодируемого плазмидой, концентрация которого в клетке тем выше, чем больше число внутриклеточных копий плазмиды. Действительно, анализ механизмов репликации мутантных плазмид, для которых характерна высокая копийность, позволил выявить два *транс*-действующих фактора, кодируемых плазмидой и оказывающих влияние на репликацию плазмиды *in vivo*.

Основным ингибитором репликации оказалась небольшая РНК длиной в 108 нуклеотидов, названная РНК I, полностью комплементарная 5'-концевой последовательности предшественника праймера (РНК II). Промотор гена РНК I расположен в области начала репликации плазмиды ColE1 и направлен в противоположную сторону по отношению к промотору РНК II (см. рис. 1.48). Комплементарные взаимодействия между РНК I и РНК II оказывают влияние на образование пространственной структуры РНК II таким образом, что предпочтительно возникает конформация $\beta\gamma$, неактивная в отношении инициации репликации (см. рис. 1.48, б, внизу справа).

Взаимодействие между РНК I и РНК II происходит продуктивно лишь до тех пор, пока синтезируется короткий транскрипт РНК II длиной не более 80 нуклеотидов. Хотя взаимодействие РНК I с такой короткой последовательностью нуклеотидов происходит медленнее, чем с транскриптом длиной в 360 нуклеотидов, в последнем случае РНК I не оказывает влияния на

конформацию 5'-концевой части РНК II и на ее способность функционировать в качестве затравки при репликации плазмиды (конформация $\alpha\beta$, рис. 1.48,б, внизу слева). Из этого ясно, что скорость образования гибридов между РНК I и РНК II является определяющей для эффективного функционирования механизма регуляции репликации плазмиды. Процесс взаимодействия РНК I и РНК II в настоящее время детально изучен. Он проходит через образование нескольких промежуточных продуктов и завершается получением стабильного гибрида между полностью комплементарными друг другу РНК I и 5'-концевой областью РНК II.

РНК-организующий белок Rop. Ген второго компонента, негативно регулирующего репликацию плазмиды ColE1, картирован непосредственно за областью начала репликации. Этот ген кодирует 63-звенный белок, названный Rop (repressor of primer), существующий в растворе в виде димера. Как *in vivo*, так и *in vitro* Rop усиливает ингибирующую активность РНК I, не оказывая влияния на синтез РНК II. При этом Rop влияет на начальные фазы взаимодействия РНК I и РНК II, облегчая переход очень нестабильного промежуточного продукта C^* в более стабильный – C^{m*} . Белок Rop обладает высоким сродством к C^* и лишь слабо взаимодействует с изолированными РНК I и РНК II *in vitro*. Предполагают, что Rop проявляет незначительную специфичность в отношении последовательностей нуклеотидов и распознает некоторые общие особенности структуры комплекса РНК I–РНК II, возникающего на ранних этапах их взаимодействия. Таким образом, функции белка Rop, по-видимому, заключаются в преобразовании нестабильного комплекса РНК–РНК в более стабильный, что, в свою очередь, сопровождается подавлением формирования праймера, необходимого для инициации репликации плазмиды ColE1.

Использование антисмысловых РНК в контроле репликации бактериальных плазмид является распространенным приемом. В частности, репликация небольшой, низкокопийной плазмиды R1 контролируется белком RepA, который участвует в инициации репликации плазмиды в качестве позитивного регуляторного фактора. Синтез RepA, в свою очередь, регулируется посттранскрипционно с помощью небольшой антисмысловой РНК SopA, которая связывается с RepA–мРНК в результате многоступенчатой реакции, напоминающей образование гибрида между РНК I и РНК II,

рассмотренное выше. Такое взаимодействие подавляет экспрессию гена *repA*, возможно, вследствие расщепления РНК–РНК-дуплекса РНКазой III. Внутриклеточная концентрация антисмысловой CopA-РНК прямо пропорциональна числу копий плазмиды R1. Аналогичный механизм описан и для регуляции инициации репликации плазмиды pT181 *Staphylococcus aureus*.

При получении бактериальных векторов для генной инженерии, многие из которых содержат область начала репликации плазмиды ColE1, с целью повышения числа их копий в бактериальных клетках часто используют ингибиторы биосинтеза белка, в частности хлорамфеникол. После обсуждения механизмов регуляции контроля репликации этой плазмиды становятся понятными принципы, на которых основан данный прием. Действительно, внесение в культуральную среду хлорамфеникола блокирует биосинтез бактериальных белков, в том числе, и белка Rop, который необходим для эффективного подавления инициации репликации плазмиды под действием РНК I. В результате нарушается контроль копийности плазмид в бактериальных клетках, и они начинают непрерывно реплицироваться, используя для этой цели предварительно синтезированные бактериальные белки.

Известно, что две фенотипически различающиеся плазмиды, использующие одинаковый механизм контроля репликации, несовместимы в одной бактериальной клетке. Клетки, содержащие две плазмиды из разных групп совместимости, в процессе размножения быстро образуют две популяции, каждая из которых содержит только один тип плазмид. Это происходит вследствие случайного выбора плазмид для репликации внутри бактериальных клеток и случайного распределения исходного пула плазмид по дочерним клеткам. Эволюционное возникновение механизма контроля репликации бактериальных плазмид с использованием антисмысловых РНК расширило возможности появления плазмид, принадлежащих к разным группам совместимости и сосуществующих в одних и тех же бактериальных клетках. Действительно, несмотря на использование одного и того же механизма, антисмысловые РНК, обладающие разными последовательностями нуклеотидов, не смогут узнавать "чужие," гетерологичные РНК-мишени. Это позволяет таким плазмидам сосуществовать в одной бактериальной клетке и создает условия для их более широкого распространения в природных популяциях микроорганизмов.

4.3. Особенности репликации линейных геномов

Кольцевые замкнутые геномы характерны для многих бактерий, их плазмид и некоторых вирусов. У подавляющего большинства других организмов геном представлен линейными молекулами ДНК в составе одной или нескольких хромосом. Размышления о механизмах репликации линейных молекул ДНК породили так называемую проблему отстающей цепи ДНК, которая в природных условиях решается весьма эффективно. Проблема заключается в том, что синтез отстающей цепи ДНК, как уже было показано выше, происходит в виде коротких фрагментов Оказаки, для инициации синтеза которых требуются РНК-затравки (рис. 1.49). После удаления затравки на конце, по крайней мере, одной из вновь синтезированных в процессе репликации молекулы ДНК образуется одноцепочечная брешь, которая не может быть заполнена ДНК-полимеразой, поскольку она не функционирует в отсутствие праймера. Вследствие этого в каждом раунде репликации должно было бы происходить укорачивание хромосом с обоих концов, что приводило бы к потере генетической информации, закодированной в концевых фрагментах ДНК. Кроме того, большие размеры молекул ДНК, заключенных в индивидуальные хромосомы, требуют специальной организации их реплицирующего аппарата. В соответствии с этим представляется целесообразным кратко рассмотреть особенности репликации ДНК линейных геномов.

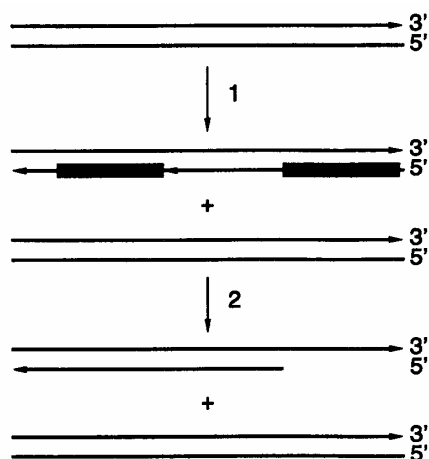


Рис. 1.49. Проблема репликации отстающей цепи ДНК линейных хромосом

1 – отстающая цепь реплицирующейся ДНК синтезируется с использованием фрагментов Оказаки, содержащих на 5'-концах РНК-затравки (зачерненные прямоугольники); 2 – по завершении синтеза затравки удаляются и фрагменты лигируются с образованием бреши на 5'-конце вновь синтезированной цепи ДНК, прилежащем к концу хромосомы

4.3.1. Линейные хромосомы бактерий

Афоризм Жака Моно: "То, что верно для *E. coli*, – верно и для других бактерий (слона)" получил широкое распространение. К счастью, на деле все обстоит не так скучно. До недавнего времени общепринятым было представление о кольцевой структуре бактериальных хромосом. Однако в 1989 г. была впервые описана у спирохеты *Borrelia burgdorferi* линейная бактериальная хромосома, которую идентифицировали с помощью электрофореза в импульсном электрическом поле. Размер этого генома составлял всего 960 т.п.о. Вскоре было обнаружено, что линейная и кольцевая хромосомы сосуществуют одновременно у *Agrobacterium tumefaciens*, а у грамположительных бактерий рода *Streptomyces*, обладающих одним из самых больших бактериальных геномов (~8000 т.п.о.), имеется одна линейная хромосома. Представитель актиномицетов *Rhodococcus fascians* также, по-видимому, обладает линейной хромосомой. Линейные хромосомы у бактерий часто сосуществуют с линейными плазмидами и широко распространены в природе.

Линейные хромосомы и плазмиды наиболее хорошо изученных бактерий рода *Streptomyces* содержат концевые инвертированные повторы (terminal inverted repeats – TIRs), с которыми ковалентно связаны концевые белки (TP). Несмотря на то что подобные структуры характерны для хромосом аденовирусов и бактериофага $\phi 29$ *Bacillus subtilis*, механизм репликации хромосом стрептомицетов существенно отличается от такового вирусных геномов. Если у вирусов синтез ДНК инициируется на конце хромосомы с использованием в качестве затравки TP, ковалентно связанного с нуклеотидом, и продолжается через весь геном до его конца, то репликация хромосомы и линейных плазмид стрептомицетов начинается с внутренней области начала репликации *oriC*. Синтез ДНК распространяется в обе стороны от области начала репликации по стандартному полуконсервативному механизму и

завершается на концах линейных молекул ДНК с образованием 3'-концевых брешей (рис. 1.50,а). Наиболее простым решением проблемы заполнения этой брешки могла бы быть прямая инициация репликации теломерных участков хромосом с ТР-белка, ковалентно связанного с иницирующим нуклеотидом, что имеет место у аденовирусов (см. рис. 1.50,б). Действительно, стрептомицеты используют ТР для репликации теломерных участков, однако механизм распознавания теломер в данном случае существенно отличается. В настоящее время рассматриваются три модели заполнения брешей в теломерных участках линейных хромосом бактерий.

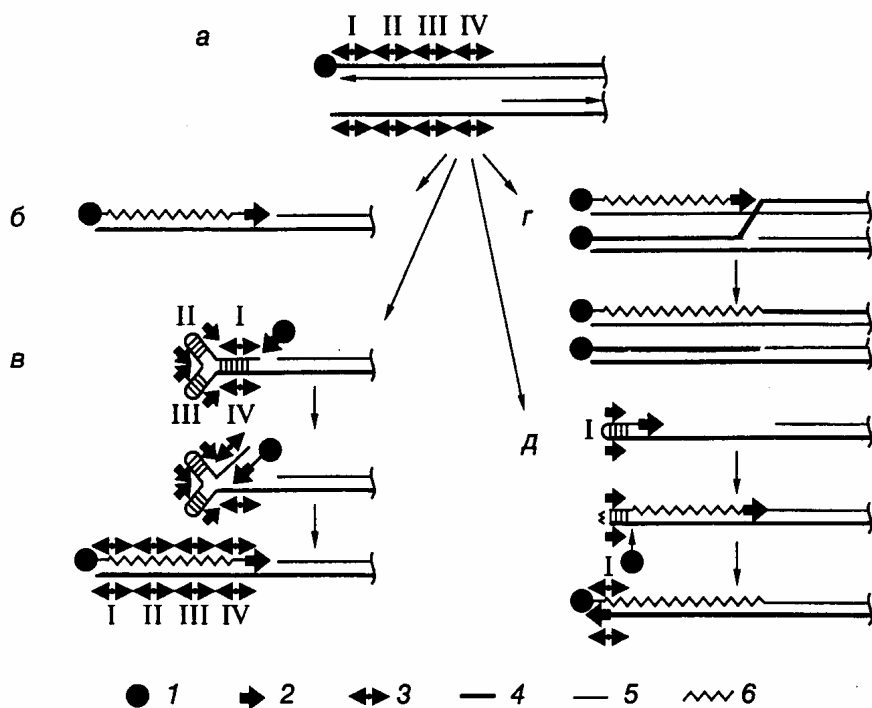


Рис. 1.50. Модель достройки теломерных участков хромосом и плазмид *Streptomyces*

а – структура теломеры после репликации: верхняя цепь ДНК полностью реплицирована, в нижней имеется одноцепочечная брешь, обозначены четыре палиндромные последовательности нуклеотидов; б – маловероятный механизм с участием концевой белка и ДНК-полимеразы; в–д – альтернативные модели репликации, основанные на других механизмах. 1 – концевой белок, 2 – ДНК-полимераза, 3 – палиндром, 4 – родительская цепь ДНК, 5 – дочерняя цепь, 6 – репаративный синтез

В соответствии с первой моделью одноцепочечный участок теломеры, содержащий TIR-последовательность, образует концевую шпильку путем

комплементарных взаимодействий нуклеотидов внутренних участков бреши и 3'-концевых нуклеотидов (см. рис. 1.50,в). В этом случае синтез ДНК, репарирующий одноцепочечную брешь, инициируется на двухцепочечном участке, образованном палиндромными последовательностями I-IV, с участием ТР и ДНК-полимеразы и продолжается вдоль 3'-концевого одноцепочечного участка хромосомы. Согласно второй модели ТР инициирует репликацию на полностью двухцепочечной дочерней ДНК, вытесняя 5'-концевую цепь родительской ДНК, с которой связан ТР (см. рис. 1.50,г). Вытесняемая цепь далее спаривается с выступающим 3'-концом хромосомы, после чего такая разветвленная структура разрешается с помощью гомологичной рекомбинации. Эта модель предполагает участие в заполнении брешей белка RecA (для переноса цепи ДНК) и продуктов генов *huv* (для разрешения структуры Холидея), что подтверждается генетическими данными. В третьей модели одноцепочечный палиндром I образует шпильку, 3'-конец которой служит затравкой для синтеза ДНК, в результате которого заполняется брешь (см. рис. 1.50,д). ТР образует одноцепочечный разрыв напротив первоначального 3'-конца, который является затравкой для последующего синтеза ДНК. В результате шпилька разворачивается и восстанавливается структура теломеры. Эта модель аналогична модели "катыщей шпильки", предложенной для объяснения механизма репликации генома парвовирусов. В данной модели роль ТР отличается от его функций в качестве белка-затравки в рассмотренных выше примерах.

Неизвестно, как много форм линейных бактериальных хромосом существует в природе. Не изучены и таксономические проблемы, связанные с топологией хромосом в царстве эубактерий. Если каждый тип хромосом характерен для отдельного таксономического домена, то можно предполагать, что топология хромосом играет важную роль в эволюции бактерий. Альтернативно топологические взаимопревращения хромосом могут быть относительно частыми событиями, а линейные и кольцевые хромосомы присутствуют только у близких видов бактерий. Нестабильность хромосом стрептомицетов (образование протяженных делеций и амплификация последовательностей нуклеотидов) недавно стали связывать с перестройками в их концевых участках, часть из которых сопровождалась образованием кольцевых хромосом. Таким образом, эволюционная роль топологии

бактериальных хромосом может быть определена только в результате будущих исследований.

4.3.2. Репликаторы эукариот

Хромосомы эукариот содержат линейные молекулы ДНК, а следовательно, остаются все те же проблемы, связанные с их репликацией, которые обсуждались в связи с воспроизводством линейных хромосом бактерий. Однако проблемы, которые эукариотическим клеткам необходимо решить при редупликации своих хромосом, несомненно серьезнее, так как размер заключенной в них ДНК значительно превышает размеры хромосомных ДНК бактериальных клеток. Кроме того, в связи с многоклеточностью большинства эукариот возникает необходимость более тонкой координации репликации ДНК в отдельных полностью дифференцированных и дифференцирующихся клетках, что является одной из основных целей регуляции клеточного цикла у данных организмов. В связи с этим организация репликации ДНК у эукариот характеризуется рядом существенных особенностей.

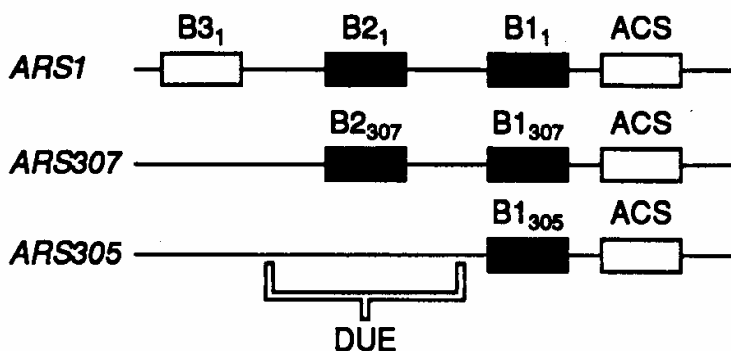


Рис. 1.51. Структура репликаторов дрожжей *S. cerevisiae*

Обозначено взаимное расположение различных регуляторных элементов в репликаторах ARS1, ARS307 и ARS305. ACS – каноническая последовательность ARS, DUE – ДНК-расплетающий элемент. Подстрочные индексы указывают на принадлежность регуляторных элементов соответствующим репликаторам

Инициация репликации у эукариот происходит на специфических множественных последовательностях нуклеотидов – репликаторах. Наиболее изученными являются репликаторы дрожжей *S. cerevisiae*, впервые

идентифицированные как автономно реплицирующиеся последовательности (ARS – autonomously replicating sequence), способные поддерживать внехромосомную репликацию плазмид в дрожжевых клетках. Исследование структуры *ARS1* показало, что этот хромосомный элемент состоит из нескольких коротких регуляторных последовательностей. Аналогичная организация характерна и для других ARS дрожжей (рис. 1.51). В частности, *ARS307* в дополнение к канонической последовательности ACS, общей для всех ARS, содержат еще два элемента – B1 и B2, которые необходимы для выполнения репликатором своих функций *in vivo*. Несмотря на то что эти последовательности в разных репликаторах не строго консервативны, внутри групп (B1, B2 и т.п.) они функционально взаимозаменяемы. Изменение положения по отношению к ACS предотвращает их функционирование.

Первым этапом инициации репликации у дрожжей является взаимодействие регуляторных последовательностей репликатора, по крайней мере, с шестью различными белками, которые образуют комплекс, распознающий область начала репликации ORC (origin-recognition complex). ARS определяет место инициации репликации в клетках дрожжей. Элемент B3 *ARS1* взаимодействует с белком Abf1, который стимулирует репликацию доменом, характерным для белков-активаторов транскрипции, тогда как B1 взаимодействует с ORC. Остающиеся регуляторные последовательности области начала репликации дрожжей образуют ранее неизвестный элемент, названный *ДНК-расплетающим элементом DUE* (DNA-unwinding element), который, как полагают, облегчает раскручивание цепей ДНК при инициации репликации. Точковые мутации в элементе B2 не влияют на функции репликатора, что является общим свойством структурных элементов, тогда как мутации в ACS, B1 и B3 нарушают инициацию репликации, как и следовало ожидать от регуляторных элементов нуклеиновых кислот, взаимодействующих с белками.

Исследования репликаторов у дрожжей *S. pombe* показали, что область начала репликации *ura4* включает в себя три отдельных репликатора, которые располагаются на участке ДНК длиной в 5 т.п.о. У млекопитающих области начала репликации располагаются на расстоянии ~100 т.п.о. друг от друга; часть их уже удалось клонировать и изучить на молекулярном уровне. Установлено, что синтез ДНК в отдельных репликонах происходит по двум

направлениям, причем перемещение репликативной вилки осуществляется предпочтительно в одном направлении, которое может изменяться в зависимости от стадии развития организма и уровня экспрессии генов, содержащих репликаторы. Частота использования отдельных репликаторов изменяется в онтогенезе, уменьшаясь в клетках взрослого организма. Сравнение первичных структур шести отдельных репликаторов эукариот показало, что все они содержат *DUE*-элементы, участки прикрепления к ядерному матриксу (SAR/MAR), канонические *ARS*-последовательности дрожжей, пиримидиновые тракты, а также ранее неидентифицированную каноническую последовательность WAWTTDDWWWHDHWGWHMAWTT, где W = A/T, D = A/C/T, H = A/C/T, а M = A/C. Имеются отдельные сообщения о том, что в репликаторах животных присутствуют пуриновые тракты, канонические последовательности, взаимодействующие с факторами транскрипции и белками репликативного комплекса, энхансерный октамерный мотив, сайты связывания продуктов онкогенов, АТ-богатые последовательности и участки перегибов (bent) ДНК. В настоящее время до конца непонятно, какое непосредственное отношение имеют все эти регуляторные последовательности к инициации репликации ДНК. Предполагается, что многие из них участвуют в регуляции транскрипции (а следовательно, и регуляции экспрессии генов) как таковой, поскольку большинство из известных в настоящее время репликаторов локализованы в 5'-концевых последовательностях функционирующих генов.

4.3.3. Репликация теломерных участков эукариотических хромосом

Исследование механизмов репликации теломерных участков эукариотических хромосом показало, что они принципиально отличаются от механизмов репликации центральных областей ДНК. Изучение этих механизмов длительное время сдерживалось тем, что в клетках животных количество хромосом, а следовательно, и теломерных участков, невелико – они составляют лишь небольшую часть от всех остальных последовательностей. Прорыв произошел в 1978 г., когда в качестве объекта исследований Е. Блэкберн стала использовать простейшее жгутиковое – *Tetrahymena*, в клетках которого содержатся тысячи линейных хромосом. Было установлено, что теломеры тетрахимены построены из коротких повторяющихся

последовательностей. Вскоре стало ясно, что такая структура характерна для теломерных участков хромосом всех исследованных эукариот. В частности, у человека теломеры содержат единственный повтор GGGTTA. Длина ДНК в теломерах хромосом человека варьирует и в клетках зародышевой линии составляет 10–15 т.п.о, а в лейкоцитах периферической крови – 5–12 т.п.о. У дрожжей длина теломер приближается к 300 п.о., и они составлены неидентичными повторами $C_{1-3}A/TG_{1-3}$. Таким образом, для теломерных участков хромосом характерна значительная гетерогенность в разных клетках и тканях даже одного организма. Еще более разительными оказываются межвидовые различия в размерах теломер – от ~50 п.о. в клетках жгутиковых до 50 т.п.о. у одного из видов мышей.

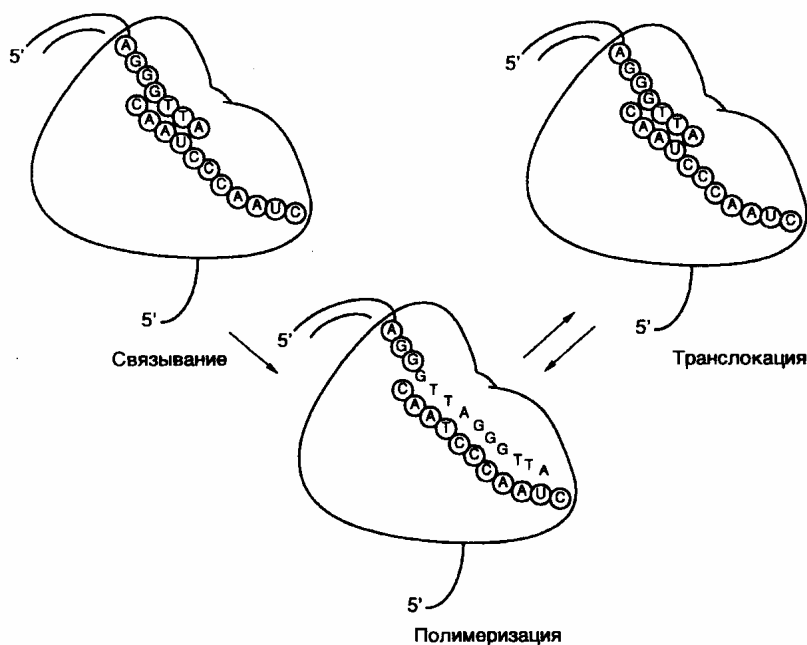


Рис. 1.52. Механизм функционирования теломеразы

Изображено комплементарное взаимодействие внутренней матричной молекулы РНК теломеразы с ДНК теломерного участка хромосомы

Синтез теломерных последовательностей ДНК осуществляется специальными ферментами – *теломеразы*, для которых нет аналогов среди других известных нуклеотидилтрансфераз. Особенностью этих ферментов является присутствие у них в качестве составной части короткого фрагмента РНК – компонента, абсолютно необходимого для их функционирования и

служащего матрицей при синтезе теломерных последовательностей хромосом (рис. 1.52). Комплементарное взаимодействие внутренней РНК теломеразы с 3'-концевым выступающим одноцепочечным сегментом ДНК хромосомы инициирует синтез теломерных последовательностей. При этом 3'-концевой фрагмент ДНК служит затравкой для удлинения этой ДНК на РНК-матрице. После удлинения (элонгации) выступающей цепи ДНК до конца матрицы происходит транслокация фермента на один теломерный повтор вперед относительно матрицы с освобождением последовательности матричных нуклеотидов, после чего он готов для вступления в следующий цикл элонгации только что добавленной 3'-концевой последовательности хромосомы. После завершения удлинения одноцепочечной 3'-концевой теломерной последовательности вторая цепь ДНК достраивается обычным способом. Таким образом происходит решение "проблемы отстающей цепи ДНК" при репликации ДНК у эукариот.

Наличие у животных тканеспецифичности в распределении теломер по размерам, а также изменение размеров этих последовательностей в онтогенезе предполагают существование механизмов, регулирующих данный процесс. Создается впечатление, что для активной пролиферации клеток теломерные последовательности не должны становиться короче определенного порогового размера. Недавние исследования обнаружили резкое повышение активности теломераз, характерное для опухолевых клеток, что служит в настоящее время чувствительным физиологическим маркером их злокачественного перерождения. В этой связи сегодня в качестве одного из подходов к терапии опухолей рассматривают подавление активности теломераз, функционирование которых, как полагают, необходимо для иммортализации клеток и роста опухолей. Именно благодаря таким свойствам теломеразы в последнее время вызывают особый интерес, что сопровождается расширением исследований в данной области молекулярной генетики.

4.3.4. Пространственная организация синтеза ДНК у эукариот

Пространственная организация репликативного синтеза ДНК у эукариот является одним из наиболее ярких примеров внутриядерной компартментализации генетических процессов. Анализ локализации мест синтеза ДНК в ядрах млекопитающих с использованием импульсной метки

бромдезоксимуридина или биотинилированного dUTP обнаруживает всего ~150 мест включения этих предшественников (центров репликации), которые приблизительно равномерно удалены друг от друга. Во время инициации синтеза ДНК размер этих центров мал, и они обнаруживаются в виде небольших четко очерченных "точек", которые со временем становятся более диффузными. В данных центрах репликации происходит аккумуляция белков, участвующих в синтезе ДНК: ДНК-полимеразы α , PCNA и RP-A, а также регуляторных молекул типа циклина A, cdk2 и RPA70. Иммунохимическими методами с использованием частиц коллоидного золота было показано, что растущие цепи ДНК выходят из центров репликации. Это позволяет предполагать, что во время репликативного синтеза цепи ДНК перемещаются через фиксированные внутри ядра структуры аппарата репликации. Такая внутриядерная компартментализация синтеза ДНК позволяет концентрировать регуляторные, структурные и ферментативные компоненты, участвующие в репликации и поддержании пространственной структуры хромосом. Ступенчатая сборка функционально активных элонгирующих комплексов в микрокомпартаментах ядра предоставляет большие возможности для регуляции инициации репликативного синтеза ДНК.

Роль ядерных пространственных структур в обеспечении функциональных свойств ДНК можно рассмотреть на примере репликации хромосом в ооцитах *Xenopus laevis*. Инъекция прокариотической ДНК в ооциты или ее добавление к экстрактам ооцитов сопровождается образованием псевдоядер, компетентных в отношении репликации ДНК. Репликация в таких системах пространственно упорядочена: она происходит в дискретных участках ДНК, содержащих кластеры репликативных вилок. При этом наблюдается замечательная корреляция между числом и пространственным распределением центров репликации в искусственных псевдоядрах и ядрах культивируемых клеток, находящихся в S-фазе. Следовательно, сборка функционирующих комплексов, способных иницировать репликацию, не находится в строгой зависимости от последовательностей нуклеотидов ДНК хромосом, но в большой степени зависит от внутренней пространственной структуры ядра и может эффективно происходить даже на прокариотических ДНК. Это означает, что пространственная структура ядра может непосредственно контролировать его функции, в данном случае – инициацию

репликации хромосом. Следует, однако, иметь в виду, что на ранних эмбриональных стадиях развития *Xenopus* контроль репликации ослаблен, и число зон начала репликации значительно больше, чем в соматических клетках. Нормальный контроль репликации устанавливается после увеличения продолжительности клеточного цикла на стадии средней бластулы.

В нормальных соматических клетках не все центры репликации одновременно начинают синтез ДНК. Некоторые из них функционируют в ранней, а некоторые в поздней S-фазе клеточного цикла. Такая дифференциальная репликация хромосом является важным регуляторным механизмом, обеспечивающим локальную организацию хроматина и активность генов. Аналогичное явление обнаружено в клетках дрожжей, где функционирование во времени центров репликации зависит от положения хромосом в ядре.

ГЛАВА 5. ЗАЩИТА ГЕНЕТИЧЕСКОЙ ИНФОРМАЦИИ

Существование биологических видов, а следовательно, и феномена жизни как такового целиком зависит от точности передачи генетической информации как по вертикали – от организмов-родителей потомкам, так и по горизонтали – от одной соматической клетки к другой в процессе онтогенетического развития многоклеточных организмов. В предыдущей главе были рассмотрены основные механизмы функционирования главной генетической системы, обеспечивающей воспроизведение генетической информации путем удвоения молекул ДНК, – системы репликации, точность функционирования которой поражает воображение. Для правильной передачи генетической информации исключительно важны и молекулярные механизмы, обеспечивающие расхождение удвоившихся хромосом между дочерними клетками у про- и эукариот, в последнем случае при митозе и мейозе. Однако ни одна из существующих генетических систем не работает безошибочно. Поэтому не менее важную роль в жизнедеятельности организма играет его *система репарации* (исправления) ошибок, случайно возникающих при репликации ДНК и после ее завершения.

У проблемы точности передачи генетической информации в ряду поколений клеток и организмов имеется и другая сторона. Чрезмерная консервация генетической информации, заключенной в отдельных генетических локусах, может быть вредной для организма и вида в целом. В частности, одним из механизмов, лежащих в основе возникновения разнообразия антител, являются запрограммированные изменения генов иммуноглобулинов, которые закрепляются в геноме лимфоцитов в результате их отбора в онтогенезе. Высокий темп изменений некоторых генетических локусов у паразитических организмов, например у трипаносом, в результате которых меняется структура антигенных детерминант на поверхности их клеток, необходим для их выживания, так как помогает этим организмам избежать нейтрализующего действия иммунной системы организма-хозяина. Другим хорошо известным примером такого рода является генетическая изменчивость вируса гриппа. Наконец, абсолютный консерватизм в передаче генетической информации по вертикали сделал бы невозможным филогенетическое развитие организмов, их

эволюционные преобразования, приведшие, в конечном счете, к тому разнообразию биологических видов, которое сегодня наблюдается в природе. Эволюционно сложившиеся отношения между точностью функционирования вышеупомянутых генетических систем и частотой ошибок, возникающих при воспроизведении генетической информации отдельных генетических локусов, четко сбалансированы между собой, и уже установлено, что в ряде случаев являются регулируемыми. Запрограммированные и случайные наследуемые изменения генома, называемые *мутациями*, могут сопровождаться колоссальными количественными и качественными изменениями в экспрессии генов.

5.1. Мутации

Мутации – это наследуемые изменения структуры генома. Поскольку основу любого генома составляют нуклеиновые кислоты – ДНК или РНК, то под действием мутаций происходит, прежде всего, изменение структуры геномных нуклеиновых кислот. Процесс возникновения мутаций, основанный на различных механизмах, называют *мутагенезом*. В зависимости от факторов, вызывающих мутации, последние принято разделять на *спонтанные* и *индуцированные*. Считается, что спонтанные мутации возникают самопроизвольно на протяжении всей жизни организма в нормальных для него условиях окружающей среды. При этом широкое распространение получило мнение о том, что спонтанные мутации в эукариотических клетках возникают с частотой 10^{-9} – 10^{-12} на нуклеотид за клеточную генерацию. Теперь, однако, становится ясно, что такие цифры не отражают реальности. Они не учитывают того факта, что частоты спонтанных мутаций могут существенно (на несколько порядков) изменяться от локуса к локусу и, скорее всего, указывают на нижний предел частоты мутаций в отдельных наиболее стабильных участках генома.

Индукцированными мутациями называют наследуемые изменения генома, возникающие в результате тех или иных *мутагенных воздействий* в искусственных (экспериментальных) условиях или при неблагоприятных воздействиях окружающей среды. Среди важнейших мутагенных факторов, прежде всего, необходимо отметить *химические мутагены* – органические и неорганические вещества, вызывающие мутации, а также *ионизирующее излучение*. При детальном рассмотрении спонтанных и индуцированных

мутаций становится ясно, что между этими двумя типами нет существенных различий. Действительно, большинство спонтанных мутаций возникает в результате мутагенного воздействия, которое их индуцирует, но не регистрируется экспериментатором. На более прочном фундаменте находится классификация мутаций, в которой учитываются молекулярные процессы, лежащие в основе их возникновения.

В классификации, основанной на размерах сегментов генома, подвергающихся преобразованиям, мутации разделяют на *геномные*, *хромосомные* и *генные*. При геномных мутациях у организма-мутанта происходит внезапное изменение числа хромосом, кратное целому геному. Если через $2n$ обозначить число хромосом в исходном диплоидном геноме, то в результате геномной мутации, называемой *полиплоидизацией*, происходит образование полиплоидных организмов, геном которых представлен $4n$, $6n$ и т.д. хромосомами. В зависимости от происхождения хромосом в полиплоидах различают *аллополиплоидию*, в результате которой происходит объединение при гибридизации целых неродственных геномов, и *аутополиплоидию*, для которой характерно адекватное увеличение числа хромосом собственного генома, кратное $2n$.

При хромосомных мутациях происходят как изменение числа отдельных хромосом в геноме (*анеуплоидия*), так и крупные перестройки структуры отдельных хромосом. Последние получили название *хромосомных аберраций*. В этом случае наблюдаются потеря (*делеции*) или удвоение части (*дупликации*) генетического материала одной или нескольких хромосом, изменение ориентации сегментов хромосом в отдельных хромосомах (*инверсии*), а также перенос части генетического материала с одной хромосомы на другую (*транслокации*) (крайний случай – объединение целых хромосом).

На генном уровне изменения первичной структуры ДНК генов под действием мутаций менее значительны, чем при хромосомных мутациях, однако генные мутации встречаются более часто. В результате генных мутаций происходят замены, делеции и вставки одного или нескольких нуклеотидов, транслокации, дупликации и инверсии различных частей гена. В том случае, когда под действием мутации изменяется лишь один нуклеотид, говорят о *точковых мутациях*. Поскольку в состав ДНК входят азотистые основания только двух типов – пурины и пиримидины, все точковые мутации с заменой

оснований разделяют на два класса: *транзиции* (замена пурина на пурин или пиримидина на пиримидин) и *трансверсии* (замена пурина на пиримидин или наоборот). Из-за вырожденности генетического кода могут быть три генетических последствия точковых мутаций: сохранение смысла кодона (*синонимическая замена нуклеотида*), изменение смысла кодона, приводящее к замене аминокислоты в соответствующем месте полипептидной цепи (*миссенс-мутация*) или образование бессмысленного кодона с преждевременной терминацией (*нонсенс-мутация*). В генетическом коде имеются три бессмысленных кодона: *амбер* – UAG, *охр* – UAA и *опал* – UGA. В соответствии с этим получают название и мутации, приводящие к образованию бессмысленных триплетов (например амбер-мутация).

По влиянию на экспрессию генов мутации разделяют на две категории: мутации типа *замен* пар оснований и типа *сдвига рамки считывания* (frameshift). Последние представляют собой делеции или вставки нуклеотидов, число которых не кратно трем, что связано с триплетностью генетического кода. Первичную мутацию иногда называют *прямой мутацией*, а мутацию, восстанавливающую исходную структуру гена, – *обратной мутацией*, или *реверсией*. Возврат к исходному фенотипу у мутантного организма вследствие восстановления функции мутантного гена нередко происходит не за счет истинной реверсии, а вследствие мутации в другой части того же самого гена или даже другого неаллельного гена. В этом случае возвратную мутацию называют *супрессорной*. Генетические механизмы, благодаря которым происходит супрессия мутантного фенотипа, весьма разнообразны.

5.1.1. Основные источники мутаций и методы определения мутагенной активности

В основе мутаций на молекулярном уровне лежат две основные причины: ошибки репликации и мутагенные воздействия различной природы. Ошибки репликации возникают из-за того, что точность функционирования ДНК-полимераз не является абсолютной. Поскольку выбор очередного нуклеотида для включения в растущую цепь ДНК определяется в результате взаимодействия белков и ферментов системы репликации и матричной ДНК, изменение свойств этих белков или матрицы как спонтанно, так и под действием различных модифицирующих агентов может приводить к ошибкам в

репликации ДНК и мутациям.

Ошибки репликации. Как уже обсуждалось в разделе 4.1, синтез ДНК происходит в результате последовательного ферментативного присоединения дезоксирибонуклеозидтрифосфатов, комплементарных матричной ДНК, к 3'-концевому нуклеотиду растущей цепи ДНК. Точность этого процесса определяется различиями в свободной энергии у канонических или ошибочных пар азотистых оснований ДНК, образующихся по правилам Уотсона–Крика в водных растворах. Различия составляют $\sim 1\text{--}3$ ккал/моль и обеспечивают точность репликации в пределах не более одного ошибочно включенного нуклеотида на каждые 100 нуклеотидов.

Генетические методы определения частоты ошибок репликации, возникающих в процессе синтеза ДНК *in vitro*, бывают двух типов. При одном подходе измеряют частоту прямых мутаций, приводящих к инактивации какого-либо гена, изменение которого легко определить фенотипически. Удобной системой, основанной на таком принципе, является репликативная форма ДНК бактериофага M13mp2, содержащая часть гена β -галактозидазы в виде одноцепочечной бреши, застраиваемой испытуемой ДНК-полимеразой в бесклеточной системе. Ошибки синтеза ДНК в этом случае выявляют по потере или уменьшению активности β -галактозидазы, что не отражается на жизнеспособности бактериофага, который образуется после введения такой ДНК в бактериальные клетки с помощью трансфекции. С применением этого метода обнаруживают до 200 различных замен оснований, а также делеции, мутации со сдвигом рамок считывания и сложные генные перестройки.

При другом подходе к определению точности синтеза ДНК *in vitro* измеряют частоту обратных мутаций, восстанавливающих (под действием точковой мутации) последовательность нуклеотидов гена, активность белкового продукта которого была нарушена в результате, например амбер-мутации. Этот метод обладает большой чувствительностью, позволяя определять мутации, происходящие с частотами $10^{-6}\text{--}10^{-8}$ на нуклеотид за одну генерацию. Однако он ограничен возможностью определения мутаций лишь одного конкретного типа.

Частота ошибок репликации увеличивается из-за спонтанных повреждений геномной ДНК, возникающих в результате ее депуринизации в

физиологических условиях. Депуринизация является следствием разрыва N-гликозидной связи, соединяющей в молекуле ДНК пуриновые основания с остатками дезоксирибозы. По оценкам Линдаля и Ниберга депуринизация ДНК в организме происходит с частотой $\sim 3 \cdot 10^{-11}$ на нуклеотид в секунду, что в ~ 100 раз выше частоты спонтанной потери пиримидиновых оснований ДНК в тех же условиях. Простой расчет показывает, что при такой частоте этих событий в соматической клетке человека должно происходить $\sim 10^5$ депуринизаций/день. Репликативный комплекс на ДНК-матрице взаимодействует с апуринизированным сайтом и включает в синтезируемую цепь ДНК преимущественно остатки дезоксиаденозина.

Многие бактериальные и эукариотические ДНК-полимеразы, осуществляющие репликацию ДНК, обладают 3'→5'-экзонуклеазной корректирующей активностью, и поэтому ошибочно включенные нуклеотиды, некомплементарные матрице, удаляются с 3'-конца растущей цепи ДНК перед включением следующего нуклеотида в строящуюся цепь ДНК. Наличие у ДНК-полимеразных комплексов такой активности существенно повышает точность функционирования систем репликации.

Мутагенные воздействия. Усилий систем репликации становится недостаточно в стрессовых ситуациях, когда организм подвергается массивному мутагенному воздействию. Однако даже присутствие незначительного количества мутагенов в окружающей среде вызывает непрерывное накопление мутаций в геноме соматических клеток, хотя и с небольшой скоростью. Процесс накопления мутаций, избежавших коррекции системами репарации, является *кумулятивным* – к ранее существовавшим мутациям неуклонно добавляются новые, и суммарное количество мутаций в геноме (*генетический груз*) возрастает. Процесс накопления мутаций – статистический, поэтому в настоящее время можно предсказывать лишь вероятность возникновения конкретной мутации в генетическом локусе или геноме организма. Поскольку мутации часто являются причиной наследственных и приобретенных заболеваний, важно делать статистический прогноз частоты возникновения определенных заболеваний в популяциях, для которых известна мутагенная экологическая обстановка.

Ионизирующее излучение. Ярко выраженным мутагенным действием обладают коротковолновое электромагнитное излучение (УФ-свет,

рентгеновские лучи), а также элементарные частицы, образующиеся в процессе радиоактивного распада вещества. С помощью рентгеновского излучения Г. Меллером в 1927 г. впервые были получены мутации у дрозофилы. С тех пор исследования механизмов мутагенеза проводятся все интенсивнее и в широких масштабах.

Электромагнитное излучение, проходящее через вещество, или элементарные частицы передают свою энергию атомам. В результате первичного столкновения с квантами излучения или частицами из атома, который превращается в положительно заряженный ион, выбиваются электроны. Освобожденные электроны вторично вызывают образование пар ионов при перемещении до тех пор, пока их энергия не иссякнет и они не утратят свою ионизирующую способность. Единицей дозы излучения служит *рентген (Р)* – количество излучения, которое вызывает образование $2 \cdot 10^9$ пар ионов/см³ воздуха. На практике часто пользуются единицей *рад*, служащей мерой поглощения энергии; в воздухе 1 Р эквивалентен 0,876 рад. Для объяснения механизмов возникновения мутаций под действием ионизирующего излучения была применена ранее разработанная *теория мишени*, в соответствии с которой повреждение ДНК наблюдается в том месте, где имеет место первичная ионизация. Реакция происходит внутри дискретного объема, являющегося мишенью. Повреждения ДНК возникают как вследствие прямого попадания кванта излучения или элементарной частицы в молекулу, так и в результате вторичного действия иона, образовавшегося за пределами ДНК в некоем "*чувствительном объеме*". Установлено, что частота мутаций, возникающих у дрозофилы и других объектов, прямо пропорциональна дозе облучения. Определенная доза облучения вызывает одинаковое число мутаций как при однократном, так и при дробном облучении небольшими порциями.

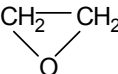
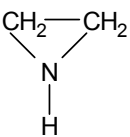
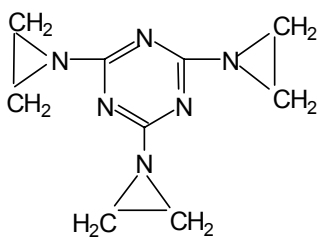
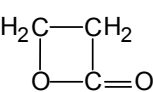
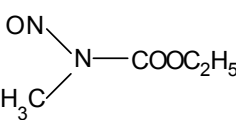
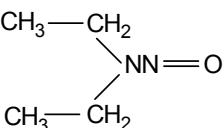
Химические мутагены экзогенного происхождения. Многие химические соединения, встречающиеся в окружающей среде, обладают способностью взаимодействовать с ДНК или с ее низкомолекулярными предшественниками и вызывать мутации. При этом одни химические соединения изначально являются реакционноспособными мутагенами, непосредственно соединяющимися с ДНК и изменяющими ее химическую структуру, а другие, так называемые *промутагены*, для превращения в мутагены сначала претерпевают *метаболическую активацию* под действием

ферментативных систем организма.

Одним из наиболее обширных классов химических мутагенов экзогенного происхождения являются *алкилирующие агенты*, под действием которых происходит спонтанный (без участия ферментативных систем организма) перенос алкильных групп этих химических соединений на биологические макромолекулы, в том числе и ДНК. В табл. I.18 представлены основные классы алкилирующих агентов. В структурах химических соединений, приведенных в таблице, можно выявить различия по двум признакам, роль которых в мутагенной активности алкилирующих агентов была неоднократно подтверждена экспериментально. Один из признаков относится к типу переносимых алкильных групп: метильной, этильной или более сложной. Другой отличительный признак – число алкильных групп, которые отдает одна молекула алкилирующего агента. Это свойство называется *функциональностью* соединения. Так, среди азотистых ипритов $\text{H}_2\text{NCH}_2\text{CH}_2\text{Cl}$ – монофункционален, $\text{HN}(\text{CH}_2\text{CH}_2\text{Cl})_2$ – бифункционален, а $\text{N}(\text{CH}_2\text{CH}_2\text{Cl})_3$ – трифункционален.

Главным источником мутаций, возникающих под действием алкилирующих агентов, является алкилирование О-6 в гуанине и О-4 в тимине ДНК. Другими сайтами, алкилирование которых реже приводит к мутациям, могут быть N-3 гуанина, N-1, N-3 и N-7 аденина, N-3 цитозина, а также N-3 и N-4 тимина. При этом спектр мутаций, возникающих под действием любого алкилирующего агента, как правило, специфичен. Следует отметить, что благодаря функционированию репаративных систем клетки к возникновению мутаций приводит лишь небольшая часть алкилирований ДНК. Поэтому частота реакций между алкилирующим агентом и ДНК не связана простой зависимостью с их мутагенной активностью. То же самое относится не только к алкилирующим агентам, но и к другим мутагенам.

Основные классы алкилирующих агентов

Класс	Представитель	Структурная формула
Иприты сернистые	Иприт	$S(CH_2CH_2Cl)_2$
азотистые	Азотистый иприт	$NN(CH_2CH_2Cl)_2$
Эпоксиды	Этиленоксид	
Этиленимины	Этиленимин	
	Триэтиленмеламин	
Алкилалкансульфонаты	Этилметансульфонат	$C_2H_5OSO_2CH_3$
	Метилметансульфонат	$CH_3OSO_2CH_3$
β -Лактоны	β -Пропиолактон	
Диазосоединения	Диазометан	$CH_2N=N$
Нитрозосоединения	N-Нитрозо-N-метилуретан	
	Диэтилнитрозамин	
	N-Метил-N-нитро-N-нитрозогуанидин	$CH_3N(NO)C(=NH)NHNO_2$

Алкилирование является одной из возможных химических модификаций нуклеотидов ДНК *in vivo* под действием экзогенных мутагенов. Число известных химических веществ, способных вызывать модификации нуклеотидов ДНК по другим механизмам, быстро возрастает с расширением исследований в этой области. Среди таких мутагенов следует упомянуть *азотистую кислоту*, которая образуется из нитритов (NaNO_2 и KNO_2) в водных растворах при низких значениях pH. Азотистая кислота дезаминирует гуанин до ксантина, аденин до гипоксантина, а цитозин до урацила. В ДНК спаривание урацила с аденином приводит к транзциям GC→AT, гипоксантин вызывает обратную транзицию AT→GC, ксантин же не спаривается ни с одним из пиримидинов ДНК, и его включение оказывается летальным для клетки. Мутагенным действием обладают различные *органические перекиси*. Сама перекись водорода не оказывает мутагенного эффекта, но становится сильным мутагеном в сочетании с формальдегидом или ацетоном, у которых индуцирует образование свободных радикалов. *Азид натрия* – мощный ингибитор дыхания, также в ряде случаев обладает мутагенным действием, что связывают с накоплением в процессе метаболизма мутагенных перекисей. Перекиси индуцируют мутации и разрывы хромосом, имитируя мутагенное действие рентгеновских лучей, которые индуцируют образование различных реакционноспособных радикалов. И, наконец, среди химических мутагенов необходимо упомянуть аналоги нуклеозидов и оснований: *5-бромдезокситуридин* и *2-аминопурин*, являющиеся сильными мутагенами. Спаривание с аденином 5-бромдезокситуридина, обычно включающегося в ДНК вместо цитозина, приводит к образованию транзиций GC→AT. Обратные транзиции AT→GC возникают под действием 2-аминопурина.

Красители, обладающие способностью интеркалировать между основаниями ДНК, вызывают мутации со сдвигом рамки считывания. К таким красителям, в частности относятся хорошо известные *бромистый этидий* и производные *акридина*.

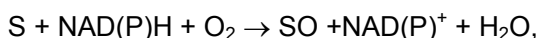
Промутагены, проканцерогены и их метаболическая активация.

Рассмотренные выше химические соединения изначально обладают мутагенной активностью и для ее проявления не требуют каких-либо дополнительных модификаций. Однако существует множество химических

соединений, которые сами по себе не проявляют мутагенную активность, но приобретают ее после серии биохимических превращений внутри организма под действием его ферментных систем. Такие химические соединения получили название промутагенов, а процесс их ферментативного превращения в мутагены – *метаболической активации*.

Любой организм в течение жизни подвергается непрерывной атаке экзогенными чужеродными химическими соединениями – *ксенобиотиками*, для защиты от накопления которых у него имеются эффективные ферментативные механизмы. Процесс метаболизма ксенобиотиков в организме традиционно разделяют на две фазы. Поскольку многие опасные для здоровья ксенобиотики гидрофобны и имеют тенденцию к накоплению в липидах клеточных мембран, обе фазы их внутриклеточного метаболизма связаны, прежде всего, с введением в их молекулы гидрофильных химических групп. Последующее повышение растворимости этих соединений в биологических жидкостях облегчает их выведение из организма. Таким образом, во время первой фазы метаболизма ксенобиотиков происходит ферментативное введение в их молекулы небольших полярных групп (например гидроксильных), что делает молекулы ксенобиотиков более водорастворимыми. С другой стороны, такие реакции подготавливают молекулы ксенобиотиков для вступления во вторую фазу метаболизма, во время которой происходит конъюгация преобразованных в первой фазе молекул ксенобиотиков с еще более полярными химическими группировками, в частности с остатками глюкуроновой кислоты, сульфатов или глицина.

Среди ферментов, принимающих участие в первой фазе метаболизма ксенобиотиков, наиболее хорошо изученными являются цитохромы группы P-450. Эти ферменты в организме представлены большим числом (>20) изоформ, каждая из которых, как правило, обладает широкой субстратной специфичностью, а все вместе они способны метаболизировать множество экзогенных и эндогенных химических соединений. Обобщенное уравнение химической реакции, осуществляемой цитохромами P-450, представлено ниже:



где S – молекула субстрата, а SO – ее окисленная форма. Таким образом, для

окисления субстрата цитохром Р-450 чаще всего использует электроны свободного кислорода, поступление которых к ферменту опосредовано NAD(P)H и флавопротеиновой редуктазой. Молекулы цитохрома Р-450 интегрированы в мембраны и при гомогенизации клеток ассоциированы с фракцией микросом. На рис. 1.53 представлен механизм метаболической активации 1,2-бензпирена (БП) – одного из наиболее сильных *проканцерогенов*. БП широко распространен в окружающей среде и в больших количествах образуется при горении нефтепродуктов, табака, а также при других термических воздействиях на органические соединения. Метаболические превращения БП по *эпоксид-диольному пути* начинаются с его эпоксидирования в положении 7,8 определенной изоформой цитохрома Р-450, ассоциированной с мембранами микросом. Полученный эпоксид гидролизует эпоксидгидролазой в соответствующий диол БП. Далее под действием другой изоформы цитохрома Р-450 происходит образование предпочтительно двух диастереомеров эпоксида в положении 9,10 и других минорных производных (соединения в фигурных скобках, см. рис. 1.53). Оказалось, что изомер (+)-анти-БПДЭ как *in vitro*, так и *in vivo* нестабилен и ковалентно связывается преимущественно с N2-атомом дезоксигуанозина ДНК своим C10-атомом с образованием *аддуктов* БПДЭ-ДНК.

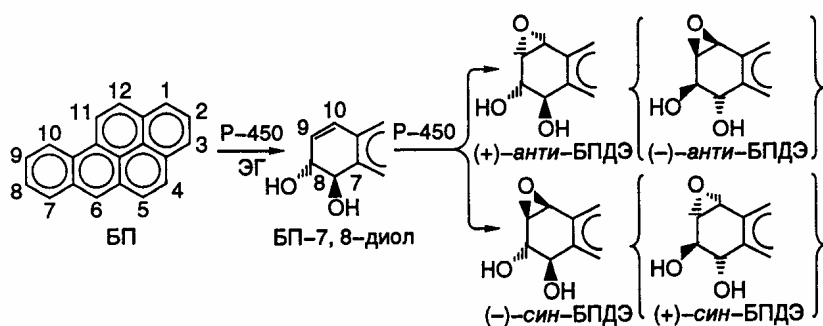


Рис. 1.53. Механизм метаболической активации 1,2-бензопирена (БП) микросомами животных

БПДЭ – 9,10-эпоксид-1,2-бензопирен-7,8-дигидродиола, Р-450 – цитохром Р-450, ЭГ – эпоксидгидролаза

Механизм метаболической активации БП подробно рассмотрен здесь для иллюстрации общих принципов химического мутагенеза. Метаболическая активация ксенобиотиков с образованием реакционноспособных метаболитов

является одной из основных причин *генотоксичности* множества химических соединений окружающей среды, с которыми организм контактирует на протяжении всей жизни. Список химических мутагенов экзогенного происхождения далеко не исчерпывается перечисленными химическими соединениями. По мере углубления исследований число известных мутагенов непрерывно растет благодаря созданию высокочувствительных методов оценки мутагенной активности. Известно, что почти все мутагены одновременно являются *канцерогенами* – веществами, вызывающими возникновение злокачественных опухолей, хотя обратное утверждение не всегда верно. Все это объясняет необходимость сохранения чистоты среды обитания человека, животных и растений.

Эндогенные мутагены. Выше уже упоминалось, что молекулы ДНК часто претерпевают *in vivo* тепловую депуринизацию, которая может быть спонтанным внутренним источником измененных нуклеотидов. Другим источником эндогенных мутаций служит самопроизвольное дезаминирование цитозина в составе ДНК с образованием урацила. 5-Метилцитозин – одно из модифицированных оснований ДНК, представляет собой "*горячую точку*" возникновения мутаций путем спонтанного дезаминирования, так как в результате удаления его аминогруппы образуется нормальное основание Т, не распознаваемое системами репарации как мутантное.

К сожалению, не только окружающая среда может быть источником генотоксических химических соединений, отравляющих жизнь организма и нарушающих функционирование его генома. Свободные радикалы, которые образуются в организме во время нормального метаболизма, могут быть дополнительной и существенной причиной спонтанных мутаций. В частности, частота повреждений нуклеотидов ДНК, появляющихся под действием *свободных радикалов кислорода*, приближается к частоте мутаций, возникающих при депуринизации ДНК. В клетках свободные радикалы кислорода возникают в реакциях восстановления, в результате которых появляются чрезвычайно реакционноспособные промежуточные соединения. Наибольшую опасность для ДНК представляют радикалы гидроксила, супероксид и синглетный кислород, которые образуются в процессе дыхания, фагоцитоза и при повреждении клеток. Путем измерения содержания в моче человека 8-гидроксидезоксигуанозина и тимингликоля – основных

модифицированных нуклеозидов и оснований, образующихся под действием радикалов кислорода, было установлено, что ежедневно в каждой клетке человека возникает ~10000 таких модифицированных нуклеотидов, мутагенный эффект которых в настоящее время доказан.

Другим источником эндогенных мутагенов в организме является метаболизм нормальной микрофлоры. Некоторые промежуточные соединения, возникающие при метаболизме аминокислот, желчных кислот и холестерина под действием бактерий организма человека и обладающие мутагенной и канцерогенной активностью, представлены в табл. I.19. *Промоторами* в канцерогенезе называют вещества, не имеющие канцерогенной активности, но ускоряющие процесс канцерогенеза под действием химических веществ – истинных канцерогенов. Как видно из таблицы, достаточно широк список веществ, обладающих всеми вышеупомянутыми активностями и образующихся под действием нормальной микрофлоры человека, обитающей, в частности в его кишечнике и на слизистых оболочках. Помимо этого в результате метаболизма в бактериальных клетках происходит активация нитратов с образованием мутагенных и канцерогенных N-нитрозаминов в кишечнике, желудке и ротовой полости человека. Описан также гидролиз обезвреженных ферментами детоксикации конъюгатов истинных канцерогенов.

Таким образом, многочисленные мутагены экзогенного и эндогенного происхождения создают вокруг и внутри любого многоклеточного организма тот мутагенный фон, к которому он вынужден приспосабливаться в процессе эволюции. Мутагенный фон "среды обитания" информационных макромолекул и, прежде всего, ДНК уникален для каждого вида живых организмов в силу особенностей их метаболизма и условий жизни. В соответствии с этим каждый биологический вид для того чтобы выжить, т.е. понизить частоту мутаций в конкретных генах до приемлемого уровня, обладает системами защиты от мутаций определенной эффективности. О механизмах функционирования таких систем защиты генетической информации речь пойдет ниже. А пока рассмотрим некоторые методы определения мутагенной активности химических соединений, применяемые для *мониторинга* окружающей среды.

Таблица I.19

Метаболиты нормальной микрофлоры человека, обладающие мутагенной и канцерогенной активностями

Соединение	Метаболит	Тип активности
Метионин	Этионин	К
Тирозин	Фенол	П
	<i>p</i> -Крезол	П
	4-Этилфенол	П
Триптофан	Индол	П
	Индолилуксусная кислота	П
	3-Гидроксикинуренин	П, М
	3-Гидроксиантраниловая кислота	П, М
	8-Гидроксихинальдиковая кислота	П
	Хинальдиковая кислота	П
	Ксантуреновая кислота	П
Желчные кислоты	Дезоксихолиевая кислота	М, П
	Литохолиевая кислота	П
	бис-Нор-5-холеновая кислота	К
	Апохолиевая кислота	К
Холестерин	Эпоксиды метаболитов: 5 β -холестерин-3 β -ола, 5 β -холестерин-3-она, 4-холестерин-3-она и др.	М, К

Примечание. К – канцероген, П – промотор, М – мутаген.

Методы определения мутагенной активности химических соединений. В результате интенсивного исследования мутагенеза было разработано несколько чувствительных методов, позволяющих с высокой эффективностью выявлять мутагенную активность химических соединений. В основном используются два подхода к определению мутагенов. При одном из них модельные организмы (бактерии или соматические клетки) выращивают в присутствии тестируемых веществ и количественно оценивают интенсивность образования клеток с мутантным фенотипом. Другая группа методов основана на прямом определении аддуктов мутагенов с макромолекулами с помощью высокоэффективной жидкостной или газовой хроматографии, а также масс-спектрометрии. Рассмотрим принципы двух биологических методов, часто используемых для определения мутагенной активности разнообразных веществ.

Тест Эймса. Благодаря своей простоте и высокой эффективности он применяется наиболее широко. Принцип метода основан на том, что мутантные бактериальные клетки, *ауксотрофы* по какому-либо метаболиту, выращивают в присутствии исследуемого вещества. Ауксотрофными называют мутантные бактериальные или иные клетки, неспособные расти на минимальной питательной среде без добавок метаболита, биосинтез которого нарушен в результате мутации. В классическом тесте Эймса чаще всего используют клетки *Salmonella typhimurium*, ауксотрофные по аминокислотам (в частности Trp), ауксотрофность которых возникает в результате единственной точковой мутации. Суспензию бактерий инкубируют в присутствии исследуемого вещества и высевает на твердую питательную среду, содержащую минимальное количество вещества, по которому бактерии ауксотрофны. Вещества должно быть достаточно для того, чтобы клетки смогли совершить несколько делений, не образовав видимых колоний. При этом образующиеся мутации фиксируются в геноме бактерий. В том случае, когда возникают реверсии в локусе, определяющем их ауксотрофность, такие бактерии приобретают способность расти без вышеупомянутых добавок на твердых питательных средах, образуя видимые колонии. Таким образом, если испытуемое вещество обладает мутагенным действием, эффективность образования ревертантов в его присутствии будет значительно выше, чем без него, что проявляется в виде множества бактериальных колоний, вырастающих

на твердой питательной среде в отсутствие пищевой добавки. Имеется большое количество модификаций теста Эймса. Наиболее важной является использование микросомных фракций печени грызунов для активации промутагенов, которые не обладают мутагенной активностью, но приобретают ее после метаболической активации в организме животных. Микросомные фракции печени содержат все основные ферменты метаболической активации ксенобиотиков, поэтому, если испытуемое вещество обладает промутагенной активностью, оно активируется ферментами микросом *in vitro*, и такой результат легко выявляется в тесте Эймса.

Сестринские хроматидные обмены (СХО). Генотоксическое действие химических веществ *in vivo* часто сопровождается дестабилизацией генома и его перестройками. Если контакт с мутагенами приводит к образованию разрывов ДНК, в процессе их репарации наблюдают обмены гомологичными участками сестринских хроматид в интерфазных ядрах, частота которых является мерой мутагенного воздействия на клетки. Разработаны эффективные методы, позволяющие обнаруживать СХО в соматических клетках животных и растений. Эти клетки (обычно лимфоциты периферической крови) инкубируют *in vitro* в присутствии 5-бромдезоксимуридина, который включается в их ДНК вместо тимидина на протяжении двух клеточных циклов. При этом одна из хроматид в хромосомах включает аналог нуклеозида в обе цепи ДНК, тогда как другая – только в одну. Флуоресцентный краситель по-разному взаимодействует с такими гомологичными хроматидами, что выявляется с помощью обычного микроскопа по различной интенсивности их окраски – одна из гомологичных хроматид выглядит светлее другой. При наличии СХО светлые и темные участки в хроматидах индивидуальных хромосом чередуются, и число таких перемежающихся участков является мерой частоты СХО. Чем выше СХО, тем интенсивнее было предшествующее мутагенное воздействие на соматическую клетку. В отличие от теста Эймса определение СХО может быть использовано для ретроспективной оценки мутагенного воздействия на организм человека и животных, поскольку СХО продолжают некоторое время после иницилирующего действия мутагена уже в его отсутствие, например после устранения источника ионизирующего излучения или полной детоксикации ксенобиотика.

5.1.2. SOS-мутагенез у бактерий

Образование мутаций в клетках организма, подвергнутого мутагенному воздействию, происходит в основном по одному и тому же механизму. При прохождении репликативного комплекса через некодирующий или ошибочно кодирующий поврежденный участок ДНК наблюдается включение в синтезируемую цепь случайных или соответствующих мутантному участку нуклеотидов. Затем репликация ДНК продолжается в обычном режиме. Таким образом, природа первичных повреждений матричной ДНК определяет тип мутационных замен нуклеотидов в ДНК.

У *E. coli* мутагенез под действием УФ-света и многих химических веществ происходит в результате координированной экспрессии большого числа генов, индуцируемых в ответ на повреждение ДНК. Такая реакция бактериальных клеток на генотоксические воздействия получила название *SOS-ответа*, а сам процесс образования мутаций – *SOS-мутагенеза*. В индукции SOS-ответа у *E. coli* определяющую роль играют два гена: *lexA* и *recA*. Белок LexA является репрессором гена *recA* и более 20 других генов и оперонов, составляющих *SOS-регулон*. В ответ на повреждение ДНК или ингибирование репликации, как правило, при прохождении ДНК-полимеразой поврежденного участка ДНК вырабатывается внутриклеточный SOS-сигнал. При этом продукт гена *recA* связывается с одноцепочечными участками ДНК и в результате конформационного перехода обратимо превращается в активированную форму RecA*. Молекулы белка LexA диффундируют к RecA* и взаимодействуют с ним, что сопровождается протеолитическим расщеплением LexA вблизи середины его полипептидной цепи (связь Ala-84–Gly-85), что инактивирует LexA как репрессор. В результате происходит индукция LexA-зависимых генов SOS-регулона. В этой реакции белок RecA* не действует как протеиназа, но активирует криптическую протеиназную активность LexA, что завершается расщеплением полипептидной цепи репрессора по аутокаталитическому механизму.

Исследование мутантов *E. coli*, неспособных к SOS-ответу при УФ-повреждениях, привело к открытию еще одного ключевого локуса *umuC*. Оказалось, что этот локус является опероном, состоящим из двух генов – *umuD* и *umuC*. Потеря функции любого из данных генов приводит к подавлению SOS-

ответа мутантных бактерий. Таким образом, *umuCD*-оперон, находящийся под контролем репрессора LexA, входит в состав SOS-регулона, и его функционирование абсолютно необходимо для SOS-мутагенеза. Во время SOS-ответа белок UmuD также расщепляется по аутокаталитическому механизму, запускаемому RecA*. Образующийся в результате полипептид UmuD' с молекулярной массой 12 кДа, включающий С-концевые остатки UmuD, оказался одним из самых важных компонентов системы SOS-ответа и SOS-мутагенеза у *E. coli*. Исследования различных мутантных производных UmuD показали, что его нативная форма не является просто неактивным предшественником белка UmuD', а выполняет функции репрессора SOS-ответа. В растворе UmuD и UmuD' существуют как в виде гомодимеров, так и более стабильного гетеродимера UmuD–UmuD', каждый из которых может объединяться с белком UmuC. Комплекс (UmuD')₂–UmuC запускает SOS-ответ, тогда как тримеры (UmuD)₂–UmuC и UmuD–UmuD'–UmuC являются в этом отношении неактивными. Образование последнего гетеротримерного комплекса играет важную роль в прекращении клеткой SOS-ответа, так как в этом случае активные внутриклеточные UmuD' и UmuC выводятся из реакции.

Белок RecA, кроме двух вышеупомянутых функций в индукции аутокаталитического расщепления LexA и UmuD, выполняет и третью функцию в SOS-ответе: участвует в формировании так называемых нуклеиново-белковых филаментов в местах одноцепочечных брешей на поврежденной ДНК. С этими филаментами могут соединяться белки UmuD' и UmuC в составе активного комплекса. Предполагают, что в результате такого взаимодействия комплекс (UmuD')₂–UmuC осуществляет переключение репаративного синтеза ДНК с нужд гомологичной рекомбинации на SOS-мутагенез. Среди других белковых компонентов системы SOS-мутагенеза *E. coli* необходимо отметить холофермент ДНК-полимеразы III, который проводит включение нуклеотидов в строящуюся цепь ДНК на поврежденном участке. Кроме того, для нормального функционирования системы SOS-мутагенеза необходимы продукты генов *groEL* и *groES*, являющиеся молекулярными шаперонами. Их роль, как полагают, сводится к стабилизации и обеспечению правильного фолдинга белка UmuC.

Таким образом, синтез ДНК на поврежденном участке требует наличия белков UmuD', UmuC, RecA и холофермента ДНК-полимеразы III. Это было окончательно подтверждено при использовании бесклеточной системы,

содержащей все вышеупомянутые белки, а также фрагмент одноцепочечной ДНК с единственным мутантным сайтом без одного азотистого основания. Способность ДНК-полимеразы III преодолевать поврежденный участок ДНК в отсутствие UmuD', UmuC или RecA была оценена в ~0,5%, однако при наличии всех компонентов в реакционной смеси эффективность преодоления поврежденного участка увеличивалась в 10 раз. Очищенная ДНК-полимераза I не заменяет ДНК-полимеразу III в этих опытах. Настоящая роль белкового комплекса (UmuD')₂-UmuC в данном процессе неизвестна. Предполагают, что комплекс может изменять процессивность ДНК-полимеразы, подавлять ее корректирующую 3'→5'-экзонуклеазную активность или изменять конформацию фермента на такую, при которой она начинает неадекватно оценивать пространственную структуру комплекса, образующегося с участием Уотсон–Криковских водородных связей между нуклеотидом матрицы и очередным входящим нуклеотидом. Любое из этих изменений должно сопровождаться повышением частоты ошибочного включения нуклеотидов. Недавно (1998 г.) было установлено, что тример (UmuD')₂-UmuC сам по себе обладает слабой ДНК-полимеразной активностью и, возможно, именно этот комплекс осуществляет синтез ДНК непосредственно в поврежденном участке в присутствии всех вышеупомянутых компонентов. В этой связи тример получил название ДНК-полимеразы V *E. coli*.

Белок Rev1 дрожжей, гомологичный белку UmuC *E. coli*, в очищенном состоянии обладает способностью неспецифически включать остатки dCMP в строящуюся цепь ДНК на участке матрицы, в котором отсутствуют азотистые основания. На этом основании он был отнесен к ДНК-полимеразам и был назван дезоксицитидилтрансферазой.

Уникальными свойствами обладает новая ДНК-полимераза IV *E. coli*, кодируемая геном *dinB*, которая также участвует в SOS-ответе бактерий. Очищенная до состояния, близкого к гомогенному (1999 г.), она не обладает 3'→5'-экзонуклеазной активностью и способна включать нуклеотиды в строящуюся цепь ДНК по высокодистрибутивному механизму. При каждом контакте фермента с субстратом и гибридом праймер–матрица к праймеру присоединяется единственный нуклеотид. В том случае, если вблизи 3'-конца праймера присутствует ошибочно спаренный нуклеотид (особенно пара G–G), то ДНК-полимераза IV в процессе синтеза ДНК осуществляет сдвиг рамки

считывания в результате делеции одного нуклеотида. Такой механизм может потенциально исправить мутагенные последствия смещения 3'-конца праймера по отношению к правильной рамке считывания в поврежденной ДНК-матрице. Полагают, что если во время репликации участков ДНК с простыми повторяющимися последовательностями в результате их повреждения происходит включение неправильно спаренного нуклеотида с последующим проскальзыванием 3'-конца строящейся цепи ДНК и образованием мутации со сдвигом рамки считывания, то происходит задержка репликативного комплекса и его диссоциация. В этих условиях синтез ДНК может быть продолжен ДНК-полимеразой IV, которая путем внесения дополнительных мутаций может исправить первоначальный сдвиг рамки.

Таким образом, SOS-мутагенез дает возможность микроорганизмам преодолевать летальное действие повреждений ДНК, что особенно важно при нарушениях ее структуры, которые блокируют репликацию ДНК и с которыми не справляется обычная репаративная система. В этом смысле наиболее опасны одноцепочечные бреши, в которых сохранившаяся цепь не содержит азотистых оснований. При таком развитии событий SOS-мутагенез является вторичным процессом – следствием заполнения бреши случайными нуклеотидами. Однако при некоторых видах повреждений во время SOS-ответа в растущую цепь ДНК включаются нуклеотиды, восстанавливающие ее исходную первичную структуру, т.е. происходит истинная репарация повреждений. В этой связи второй функцией SOS-мутагенеза может быть предоставление микроорганизму возможности противостоять определенным генотоксическим воздействиям окружающей среды, а именно таким классам химических мутагенов, последствия действия которых не преодолеваются системами эксцизионной или иной репарации, функционирующими с высокой точностью. Система SOS-мутагенеза, по-видимому, обеспечивает протекание последовательности реакций, обозначаемых как *"повторный старт репликации"*, что имеет место после временной (на 30–45 мин) задержки синтеза ДНК в ответ на ее повреждения генотоксическими агентами. Не исключено также, что SOS-мутагенез создает бактериальным клеткам определенные эволюционные преимущества, так как способствует поддержанию генетического разнообразия в популяциях этих микроорганизмов.

5.1.3. Мутаторный фенотип

Несмотря на обилие эндогенных и экзогенных мутагенов, лишь небольшая часть их взаимодействий с ДНК завершается образованием мутаций. Для того чтобы исходное повреждение ДНК в виде аддукта, апуринового сайта или одноцепочечного разрыва закрепилось в геноме в виде мутации, ему необходимо избежать нейтрализующего действия многочисленных ферментов системы репарации ДНК. В экспериментальных условиях для получения требуемых мутаций с помощью химических мутагенов и последующего скрининга требуется большая доза суммарного мутагенного воздействия. Из экспериментальных кривых "доза–эффект" видно, что число возникающих мутаций прямо пропорционально дозе мутагенного воздействия. Исходя из этого уже *a priori* можно предположить, что повреждение отдельных компонентов системы репарации должно приводить к возрастанию выхода мутаций в ответ на определенную дозу мутагенного воздействия. Действительно, описаны многочисленные штаммы микроорганизмов и линии соматических клеток с повышенными частотами спонтанных мутаций. Совокупность признаков организма, для которой характерна повышенная частота образования спонтанных мутаций, получила название *мутаторного фенотипа*.

Исследование молекулярно-генетических механизмов, приводящих к формированию мутаторного фенотипа, позволило обнаружить отдельные гены, ответственные за этот процесс. Такие гены называют *генами-мутаторами*, или просто *мутаторами*. Прежде всего, к ним относятся многие гены системы репарации ДНК, контролирующие разные ее этапы (подробнее см. раздел 5.3). Другие гены, мутации в которых приводят к мутаторному фенотипу, кодируют ферменты матричного синтеза нуклеиновых кислот. Описаны мутационные замены отдельных аминокислотных остатков в ДНК-полимеразах, которые понижают специфичность выбора дезоксирибонуклеозидтрифосфатов из внутриклеточного пула в соответствии с последовательностью матричной ДНК. Следствием этого является повышение частоты включения некомплементарных матрице нуклеотидов в строящиеся цепи ДНК. Однако сам процесс включения некомплементарных матрице нуклеотидов является лишь одной из стадий, критических для контроля точности репликации ДНК.

Благодаря наличию у ДНК-полимераз корректирующей 3'→5'-экзонуклеазной активности ошибочно включенные некомплементарные матрице нуклеотиды тотчас удаляются из строящейся цепи ДНК, что защищает ее от точковых мутаций, возникающих по такому механизму. Поэтому неудивительно, что мутации, нарушающие функционирование корректирующей экзонуклеазной активности, также приводят к возникновению мутаторного фенотипа. К аналогичным эффектам приводят нарушения функционирования систем рекомбинации, транскрипции, систем контроля структуры хроматина, ферментных систем, контролирующих сегрегацию хромосом и число копий индивидуальных генов, а также систем, участвующих в синтезе эндогенных мутагенов. Нарушения функционирования и координации экспрессии генов метаболизма нуклеотидов также приводят к мутаторному фенотипу. Известно, что повышение внутриклеточной концентрации дезоксирибонуклеозидтрифосфатов сверх оптимального уровня понижает точность репликации ДНК. Мутаторный фенотип после возникновения начинает имитировать непрерывное мутагенное воздействие, интенсивность которого зависит от характера повреждений генов-мутаторов.

Мутаторный фенотип у микроорганизмов проявляется в повышенной частоте возникновения спонтанных мутаций, которые можно измерить по частоте появления клеток, выживающих в селективных условиях, например в присутствии антибиотиков. У эукариот мутаторный фенотип часто сопровождается дестабилизацией генома, обнаруживаемой по возрастанию частоты внутригеномных перестроек ДНК. В связи с этим явлением наиболее интенсивно исследуются изменения структуры микросателлитных повторов в геноме человека при онкологических заболеваниях. Сравнение структуры отдельных микросателлитных локусов в клетках опухолей и нормальных тканей одного индивидуума часто обнаруживает существенные различия между микросателлитами одного и того же локуса. Такого рода исследования чаще всего проводятся с помощью ПЦР или любого другого метода, используемого при ДНК-типировании (см. главу 10).

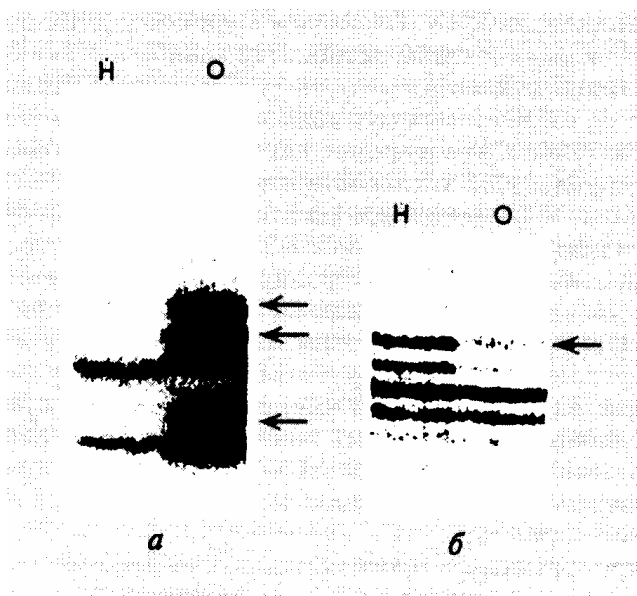


Рис. 1.54. Примеры нестабильности микросателлитов (а, на примере карциномы молочной железы) и потери гетерозиготности (б, на примере глиомы) при различных онкологических заболеваниях. Показаны продукты ПЦР после разделения с помощью электрофореза в полиакриламидном геле.

Н – нормальная ткань, О – опухоль. Стрелки указывают на измененные аллели

Наиболее просто обнаруживаются два типа изменений микросателлитов при мутаторном фенотипе (рис. 1.54). При одном из них изменение суммарной длины микросателлитных повторов конкретного генетического локуса обнаруживают по возрастанию или уменьшению электрофоретической подвижности соответствующих продуктов ПЦР при сравнении его состояния в опухолевых и нормальных тканях одного и того же организма (см. рис. 1.54, а). В исследованиях подобного рода часто наблюдают эффект так называемой *потери гетерозиготности* исследуемых микросателлитных локусов. Ввиду диплоидности генома человека каждый микросателлитный локус в нем представлен двумя копиями и, следовательно, двумя аллелями в случае его гетерозиготности. При ее потере происходит выравнивание длины обоих микросателлитных аллелей, которые различаются в нормальных тканях, или удаление одного из аллелей в результате делеции. Удаление фиксируют электрофоретически по исчезновению одного из продуктов ПЦР после амплификации соответствующих локусов, не сопровождаемому появлением

новых полос (см. рис. 1.54,б).

Дестабилизация микросателлитных локусов в опухолевых клетках, по-видимому, не служит непосредственной причиной малигнизации этих клеток. Нестабильность микросателлитов скорее может быть чувствительным маркером мутаторного фенотипа раковых клеток, внешним проявлением дестабилизации генома, характерного для опухолевых клеток многих типов. Варьирование размеров микросателлитных локусов является частным случаем большой группы мутаций, связанных с изменением числа копий последовательностей нуклеотидов в геноме эукариот. В качестве еще одного примера мутаций этого рода рассмотрим изменения размеров небольших кластеров ди- и тринуклеотидных повторов в геномной ДНК. Недавно было установлено, что такие мутации, получившие название *экспансии ДНК*, лежат в основе многих тяжелых заболеваний человека.

5.1.4. Экспансия ДНК

Под экспансией ДНК понимают увеличение числа копий коротких повторяющихся последовательностей нуклеотидов внутри кластера при передаче генетической информации от родителей потомкам. В настоящее время различают два класса этих генетических явлений. При экспансии ДНК первого класса происходит резкое и стабильное увеличение числа копий определенных повторов (≥ 10) на фоне полного отсутствия обратного сокращения длин их кластеров. При экспансии ДНК второго класса изменения затрагивают меньшее число повторов (≤ 4), а образование мутационных вставок и обратных делеций повторяющихся последовательностей происходит с одинаковой скоростью. Мутации первого класса ассоциируются с болезнью Хантингтона, синдромом ломкости X-хромосомы, болезнью Кеннеди, миотонической дистрофией, спиноцеребральной атаксией типа I и рядом других неврологических заболеваний. У здоровых индивидуумов имеет место полиморфизм по числу копий повторяющихся единиц в этих локусах, и экспансия повторов приводит к развитию таких заболеваний после того, как количество копий повторяющихся единиц (как правило, тринуклеотидов) начинает превышать определенное значение. В результате изменяется уровень экспрессии гена, содержащего повторы, или же свойства кодируемого

геном белка. Экспансия динуклеотидов СА или АТ, относящаяся ко второму классу мутаций, ассоциируется чаще всего с наследственным раком кишечника.

В наследовании предрасположенности к экспансии ДНК имеет место импринтинг (см. раздел 3.2.5). Процесс экспансии происходит во время митотической репликации ДНК на ранних стадиях эмбриогенеза. Для развития синдрома ломкости X-хромосомы *только в материнской* X-хромосоме должен образовываться премутационный аллель, содержащий 50–200 копий тринуклеотида CGG. При наличии такой премутации в эмбриональном развитии *мужской особи* формируется мутантный аллель, содержащий 200–2000 копий этого повтора. Подобной экспансии ДНК никогда не происходит, если премутация локализована *на отцовской* X-хромосоме. Приведенный пример с синдромом ломкости X-хромосомы указывает на необходимое условие для экспансии ДНК и во всех остальных случаях: увеличение числа копий повторов у родителей до определенного порогового значения. Как только формируется премутантный аллель в геноме родителей, у следующего поколения происходит образование истинно мутантного аллеля, сопровождаемое развитием соответствующего патологического состояния организма.

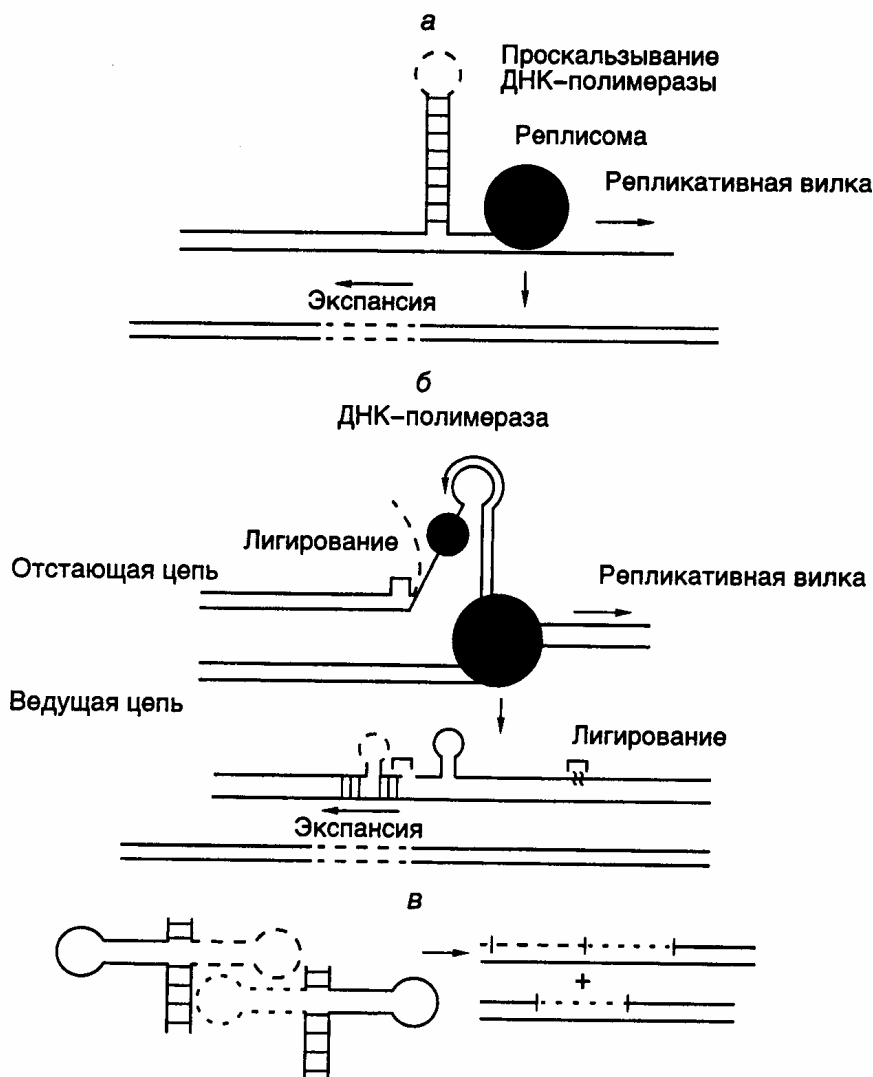


Рис. 1.55. Модели механизма экспансии ДНК во время репликации
 а – с проскальзыванием ДНК-полимеразы; б – с участием отстающей цепи ДНК; в – с привлечением механизма конверсии генов

Среди возможных ди- и тринуклеотидных последовательностей лишь для CAG/CTG, CGG/CCG и AT продемонстрирована экспансия *in vivo*, приводящая к патологиям. Предполагают, что возможность экспансии именно этих, а не других повторов определяется их способностью образовывать стабильные шпильки. Предложены три модели, объясняющие механизм экспансии ДНК, в которых участвуют шпильки, образованные соответствующими повторами (рис. 1.55). Первая модель предполагает проскальзывание реплицирующей

молекулы ДНК-полимеразы на повторе в обратном направлении, сопровождаемое образованием шпильки в строящейся цепи ДНК (см. рис. 1.55,а). В этом случае в следующем раунде репликации или по завершении репаративного синтеза ДНК длина кластера повторяющихся последовательностей увеличится на размер сегмента ДНК, образовавшего шпильку. Согласно второй модели при репликации отстающей цепи ДНК на соответствующем повторе матричная ДНК некоторое время находится в одноцепочечной форме (см. рис. 1.55,б). В это время одноцепочечная последовательность отстающей цепи, содержащая повтор, может образовать шпильку, которая блокирует синтез ДНК в данном месте. Предполагается, что белок, связывающийся с одноцепочечной ДНК (SSB-белок), не может эффективно взаимодействовать с двухцепочечной шпилькой, однако сохраняет способность связываться с одноцепочечной петлей на вершине шпильки, обеспечивая инициацию репликации ДНК с этого места. После того как ДНК-полимераза прореплицирует одну половину шпильки, ее структура разрушится, что приведет к снятию блока репликации перед шпилькой и ее нормальному продолжению. Однако синтезированный перед этим фрагмент ДНК, соответствующий одной половине шпильки, оказывается лигированным с уже имевшимся фрагментом Оказаки. Он не может отделиться от матрицы, которая своими нуклеотидами, расположенными по концам, образует водородные связи с повторами матрицы. Далее избыточная последовательность включается в растущую цепь ДНК и после ошибочной репаративной коррекции шпильки приводит к увеличению размера этого кластера повторяющихся последовательностей, т.е. их экспансии. В соответствии с третьей моделью шпильки, образующиеся в местах соответствующих кластеров повторяющихся последовательностей на разных молекулах ДНК, могут взаимодействовать друг с другом, что будет сопровождаться возникновением сложно структурированной матрицы для репликации. В процессе синтеза ДНК на такой сложной матрице весьма вероятно включение избыточной последовательности в виде повтора во вновь синтезируемую ДНК (см. рис. 1.55,в). Аналогичный эффект может быть достигнут и в результате сложной последовательности событий рекомбинации с участием этих шпилек.

Заканчивая обсуждение недавно обнаруженных "необычных" мутационных изменений геномной ДНК, связанных с увеличением числа копий

(экспансией) коротких повторяющихся последовательностей, следует подчеркнуть, что они не изменяют традиционных представлений о мутагенезе. Хотя такие мутации ассоциированы с конкретными генетическими локусами, содержащими повторы определенной структуры, процесс экспансии этих последовательностей носит случайный характер и, возможно, запускается первичным мутационным изменением в кластере повторов, которое делает более вероятным формирование шпилечной или иной пространственной структуры подобных кластеров. Однако не так давно была описана еще одна группа мутаций, само существование которых идет в разрез с общепринятым неodarвиновским представлением о механизмах возникновения мутаций. В следующем разделе речь пойдет об *адаптивных мутациях*; их возникновение в геноме носит не случайный, а направленный характер.

5.1.5. Адаптивные мутации

Проблема, связанная с возможностью возникновения адаптивных мутаций, имеет глубокие корни в биологии. За 50 лет до того как Ч. Дарвин начал свои знаменитые исследования происхождения биологических видов, другой биолог, француз Ж.Б. Ламарк, разработал учение о возможности наследования признаков, которые родительские организмы приобретали на протяжении жизни под действием окружающей среды. В терминах современной генетики это означает, что организм в ответ на определенное воздействие внешних факторов может целенаправленно изменять свой геном в клетках зародышевой линии таким образом, что он у потомков будет контролировать развитие признаков, максимально адаптированных к изменившимся внешним условиям. Следовательно, по Ламарку организм самостоятельно способен делать выбор между полезными и вредными признаками и тем самым направлять свои эволюционные преобразования. Поскольку любое наследуемое изменение генома представляет собой мутацию, неоламаркистские концепции предполагают, что организм может контролировать мутационные изменения своего генома и направлять их в нужное русло развития признаков, полезных для выживания.

Взгляды неodarвинистов на мутационный процесс прямо противоположны. В соответствии с их доктриной, доминирующей в системе взглядов современных молекулярных генетиков, мутации являются случайными

и спонтанными событиями, а образующиеся мутантные признаки подвержены жесткому давлению естественного отбора. В популяциях закрепляются лишь мутации (и признаки), максимально соответствующие условиям окружающей среды. Остальные элиминируются в результате гибели мутантных особей.

Считается, что спор между неоламаркистами и неodarвинистами окончательно решен в пользу последних, по крайней мере, тремя классическими экспериментами. Во флуктуационном тесте (С. Лурия, М. Дельбрюк, 1943 г.) исследовали возникновение мутантов *E. coli*, устойчивых к бактериофагу T1 в независимо выращиваемых культурах бактерий. Бактериальные клетки из разных пробирок высевали на чашки Петри с избытком фага T1 и подсчитывали число образующихся бактериальных колоний, устойчивых к бактериофагу. Предполагалось, что если мутантные бактерии образуются в пробирках до вступления в контакт с бактериофагом, то количество устойчивых бактерий будет сильно различаться в разных пробирках в зависимости от числа делений, которые совершит мутантная бактерия с момента возникновения мутации до посева на чашку Петри. В том случае, когда мутации возникают после взаимодействия бактерий с вирусом, количество мутантных бактерий, обнаруживаемых на разных чашках, будут следовать непрерывному распределению Пуассона. В ходе этих экспериментов продемонстрировано предсуществование мутантных бактерий в культурах. В еще более наглядных опытах Дж. Ледерберга и Е.М. Ледерберг (1950 г.) с помощью бархатного штампа перепечатывали бактерии с газона на разные чашки Петри, содержащие бактериофаг T1. Оказалось, что мутантные бактерии, устойчивые к бактериофагу, образуют колонии в одних и тех же местах разных чашек. Эти данные были также в пользу предсуществования мутантных бактерий в газоне, не имевшем контакта с бактериофагом. Однако в своих первых опытах исследователи имели дело с потомками исходных бактериальных клеток, которые контактировали с селекционирующим агентом (бактериофагом T1) в момент отбора. Поэтому "окончательное" решение вопроса было достигнуто после получения в 1956 г. штаммов бактерий, устойчивых к стрептомицину и ранее не соприкасавшихся с антибиотиком.

Потрясение твердо устоявшихся основ молекулярной генетики с неоламаркистских позиций началось в 1988 г. после опубликования в журнале "Nature" статьи Дж. Кэрнса, Дж. Овербаха и С. Миллера "Происхождение

мутантов". В серии экспериментов с мутантными клетками *E. coli*, неспособными использовать лактозу в качестве источника углерода (фенотип Lac⁻), авторы установили, что скорость образования ревертантов в том случае, если мутантные бактерии инкубировали на чашках в присутствии лактозы, значительно превышала ожидаемую из случайного возникновения обратных мутаций в стационарной бактериальной культуре. На этом основании авторы сделали вывод о том, что селективные условия (присутствие неусваиваемой лактозы в качестве единственного источника углерода) оказывают влияние на спектр мутаций, возникающих у бактериальных клеток. В работе утверждается, что бактериальные клетки могут сами контролировать свой мутационный процесс, направляя его в сторону образования нужных мутантных ферментов, что позволяет клеткам адекватно реагировать на сигналы окружающей среды, которая направленно формирует генотип бактериальных клеток.

Такие "еретические" выводы Кэрнса и его соавторов получили экспериментальное подтверждение в его дальнейших исследованиях, а также в многочисленных работах других авторов с использованием бактериальных и дрожжевых клеток в качестве объекта. И хотя в ряде случаев было показано наличие артефактов, приводивших к неправильной интерпретации результатов, в целом существование феномена направленного образования адаптивных мутаций подтверждено и пока не опровергнуто. Однако он может занять достойное место среди других хорошо доказанных генетических явлений лишь после экспериментального выяснения молекулярных механизмов, лежащих в основе адаптивных мутаций.

Для объяснения этих фактов в настоящее время выдвинуто несколько гипотез, ни одна из которых пока не получила полного экспериментального подтверждения. По мнению Кэрнса и соавторов (1988 г.) клетки синтезируют набор переменных, незначительно различающихся по первичной структуре молекул мРНК, и путем обратной транскрипции получают копию кДНК с одной из них, кодирующей наиболее подходящую для адаптации белковую молекулу. Далее такая кДНК в результате рекомбинации замещает мутантный аллель в геноме микроорганизма. Б.Д. Дэвис (1989 г.) считает, что индукция транскрипции отдельных локусов в геноме покоящихся микроорганизмов, в частности лактозой, повышает их мутабельность. Ф.У. Сталь (1988 г.) и Л. Боз (1990 г.) высказывают предположение о снижении функционирования систем

репарации ДНК у голодающих микроорганизмов, что может быть причиной повышения частоты мутаций в транскрибируемых локусах. Те же авторы предполагают, что в основе феномена направленного повышения частоты мутаций лежит *recA*-зависимая амплификация соответствующих генетических локусов, сопровождаемая корректирующим мутагенезом. Б.Г. Холл (1990 г.) для объяснения адаптивных мутаций разработал модель, в соответствии с которой в популяции голодающих микроорганизмов часть клеток находится в состоянии повышенной мутабельности. Среди этих клеток выживают лишь мутанты, максимально соответствующие требованиям окружающей среды.

По крайней мере, два результата исследований последних лет делают концепцию Холла наиболее приемлемой. Прежде всего было установлено, что реверсия мутантных бактериальных клеток к фенотипу Lac^+ в условиях голодания требует функционирования *RecBCD*-зависимой репарационной системы рекомбинации (см. раздел 5.2.3). Кроме того, в бактериальном геноме были обнаружены горячие и холодные точки, в которых образование адаптивных мутаций может происходить соответственно с высокой и низкой частотой, что объясняет отмеченную в литературе невозможность их обнаружения в некоторых генетических локусах. Поскольку на первых этапах работы репарационной системы рекомбинации происходит внесение в ДНК двухцепочечных разрывов, с которыми далее взаимодействует комплекс белков *RecBCD*, полагают, что такие разрывы ДНК инициируют процесс возникновения адаптивных мутаций.

Во время рекомбинационного обмена цепями ДНК, индуцированного двухцепочечными разрывами, происходит синтез новых цепей ДНК-полимеразой III, сопровождаемый ошибочным включением нуклеотидов. (Как уже упоминалось в разделе 5.1.2, ДНК-полимераза III является активным участником *SOS*-мутагенеза у бактерий.) Такие ошибочно включенные нуклеотиды с высокой вероятностью могут закрепляться в геноме в виде мутаций из-за ослабления эффективности функционирования системы эксцизионной репарации у бактериальных клеток, находящихся в стационарной фазе роста, и, следовательно, формирования у голодающих бактериальных клеток мутаторного фенотипа. При случайном возникновении мутации, возвращающей клетку к нормальному Lac^+ -фенотипу, мутантная бактерия выходит из стационарной фазы и начинает активно делиться. При этом

происходит восстановление обычного функционирования системы репарации.

Как можно видеть, обсуждаемая модель не оставляет места классическому неоламаркизму. При реализации такого механизма возникновение "адаптивной" мутации определяет случай, и после ее появления происходит клональное замещение исходной популяции бактерий мутантными клетками. Однако выбор самих генетических локусов, в которых могут происходить такие мутации, уже не является случайным. Он генетически детерминирован, на мой взгляд, структурой бактериального генома. В соответствии с развиваемой в разделе 5.3 концепцией альтруистичной ДНК, адаптивные мутации являются в конечном итоге естественным следствием дифференциальной защиты отдельных генетических локусов от спонтанных мутаций, определяемой пространственной структурой ДНК локусов. Как будет следовать из дальнейшего изложения, частота мутаций в первом приближении обратно пропорциональна уровню конденсации ДНК конкретного генетического локуса и определяется физической доступностью отдельных его частей химическим мутагенам и(или) ферментам системы репарации. Кроме того, индукция транскрипции этих преддетерминированных локусов или даже простое удаление регуляторных белков из промоторной зоны генов могли бы изменять их пространственную структуру и, как следствие, уровень мутабельности соответствующих участков генома.

5.1.6. Механизмы защиты генома от мутаций

Несмотря на то что иногда мутации помогают организму выжить, подавляющее большинство мутационных изменений генома нежелательно и сопровождается развитием различных патологических состояний мутантной особи или отдельной соматической клетки. Жестко действующий естественный отбор, в частности, через систему иммунного надзора элиминирует мутантные соматические клетки, опасные для существования многоклеточного организма, например предотвращая иногда развитие онкологических или аутоиммунных заболеваний. Однако к гораздо более плачевным последствиям приводит элиминация естественным отбором целой мутантной особи, так как это сопровождается непродуктивной гибелью большого числа соматических клеток и является расточительным с точки зрения энергетических затрат на их биосинтез. Генетическая информация любого организма тщательно защищена

от мутационных повреждений, что делает мутации в жизненно важных локусах генома очень редкими. Защита осуществляется на нескольких уровнях. Прежде всего, организм старается не допустить попадания химических мутагенов в жизненно важные локусы своего генома. Это достигается двумя путями. Во-первых, избыточные последовательности нуклеотидов ДНК, экранируя кодирующие последовательности нуклеотидов в геноме эукариот, принимают удар большей части химических мутагенов на себя, не допуская их попадания в такие локусы. Те же цели могут быть достигнуты за счет особой пространственной организации ДНК в конкретных участках генома (подробнее см. раздел 5.3). Во-вторых, в клетках имеются многочисленные высоко- и низкомолекулярные ловушки мутагенов, важнейшими из которых являются: маннит, энкефалины, индолы, желчные кислоты и их производные, α -токоферол, аскорбиновая кислота, тирозин, серотонин, а также ряд других соединений экзогенного и эндогенного происхождения.

К сожалению, обе системы защиты не обладают 100%-й эффективностью. То же можно сказать и о точности функционирования ферментных систем, осуществляющих воспроизведение генетической информации. Поэтому многочисленные нарушения первичной структуры ДНК неизбежны. Тем не менее, большинство первичных повреждений не превращается в мутации благодаря функционированию высокоэффективных систем репарации ДНК, состоящих из многих ферментных компонентов. Из-за исключительной важности функционирования систем защиты генетической информации в поддержании эффективной экспрессии генов ниже будут рассмотрены основные компоненты систем репарации ДНК и принципы их работы.

5.2. Репарация ДНК

Большая группа молекулярно-генетических явлений, известная в настоящее время под общим названием "репарация повреждений ДНК", была осознана как отдельный и очень важный биологический феномен лишь в конце 1950-х годов. По мнению Ф. Сталя такая задержка в развитии этого направления исследований была связана с широко распространенным мнением о том, что гены, как чрезвычайно тонко и точно организованные

биологические структуры, должны быть хорошо защищены от самой возможности биохимических повреждений, например путем упаковки в высокоэффективную защитную оболочку. В то время невозможно было представить себе ген в виде нестабильной макромолекулы, структура которой динамически изменяется на протяжении жизненного цикла организма, непрерывно отклоняясь от своего начального состояния и возвращаясь к исходной структуре в результате координированного функционирования большого числа ферментных систем.

5.2.1. Основные механизмы репарации поврежденной ДНК

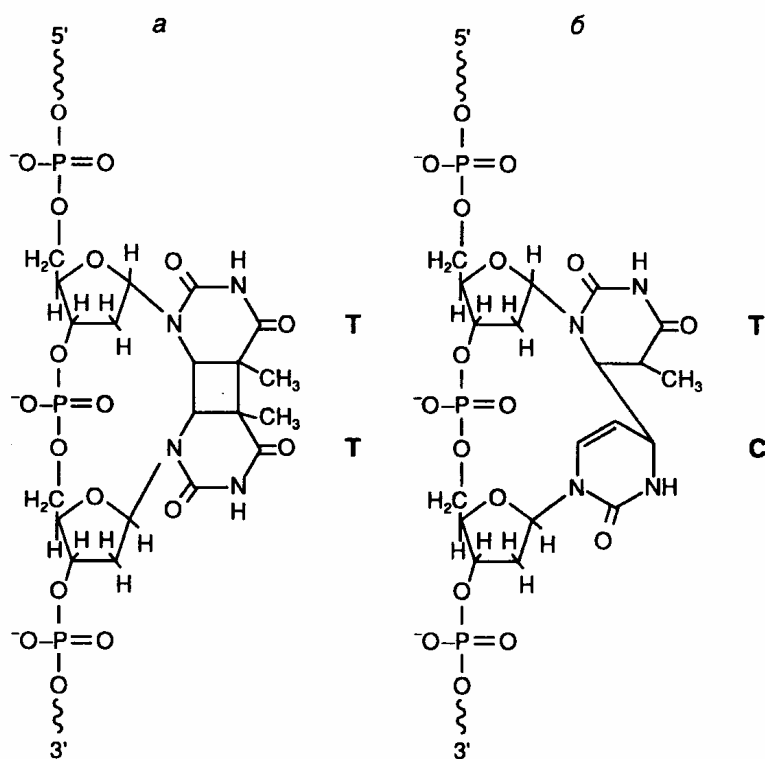


Рис. 1.56. Участок ДНК с основными повреждениями, вызываемыми УФ-светом

а – тиминовый димер циклобутанового типа; б – пиримидиновый димер, соединенный 6–4 связью.

С – цитозин; Т – тимин

Как уже упоминалось выше, имеются два типа нарушений структуры ДНК, которые в конечном итоге приводят к мутациям. Это, во-первых, включение

нормальных нуклеотидов в аномальное окружение из последовательностей нуклеотидов, приводящих к образованию неправильно спаренных оснований и петель разных размеров. Во-вторых, появление повреждений ДНК в виде аномальных нуклеотидов в правильных последовательностях ДНК. В этом случае речь идет о различных химических модификациях нуклеотидов, включая их разрушение и образование поперечных сшивок. Помимо того, что повреждения ДНК часто являются причиной мутаций, они еще могут приводить к задержке и полному блокированию репликации и транскрипции.

При исследовании механизмов репарации ДНК первые важные результаты были получены на клетках, облученных УФ-светом с длинами волн 240–280 нм. УФ-облучение клеток часто сопровождается их гибелью, образованием мутаций и злокачественной трансформацией, что вызвано в первую очередь повреждениями их ДНК. Среди первичных повреждений такого рода наиболее часто встречаются биспиримидиновые фотопродукты: пиримидиновые димеры циклобутанового типа, соединенные связью 6–4 (рис. 1.56). Как про-, так и эукариоты имеют несколько ферментных систем, которые разделяют пиримидиновые димеры или восстанавливают исходную структуру азотистых оснований. К таким репаративным системам относится, прежде всего, *система эксцизионной репарации ДНК*, осуществляющая вырезание поврежденных нуклеотидов (nucleotide excision repair – NER) или азотистых оснований (base excision repair – BER). Система ферментативной фотореактивации ДНК (photoreactivation – PHR), основным компонентом которой является *ДНК-фототлиаза*, разделяет пиримидиновые димеры, превращая их в нормальные пиримидиновые основания. Кроме того, поврежденные УФ-светом молекулы ДНК могут репарироваться с участием систем рекомбинации и в процессе пострепликативного синтеза ДНК. Действие многих вышеперечисленных систем репарации поврежденной ДНК распространяется не только на фотопродукты, но и на другие модифицированные основания, образующиеся под действием химических мутагенов. Отдельно следует упомянуть систему, распознающую неправильно спаренные основания в двойной спирали ДНК, возникающие в результате ошибок репликации.

Большинство исследованных организмов обладают системами репарации ДНК в различных комбинациях. Так, клетки *E. coli* для удаления

фотопродуктов используют системы NER и PHR, тогда как у человека пиримидиновые димеры циклобутанового типа удаляются исключительно системой NER. Системы эксцизионной репарации NER и BER благодаря своей универсальной полифункциональности занимают центральное место среди систем репарации ДНК.

5.2.2. Эксцизионная репарация в клетках животных

Эксцизионная репарация ДНК путем удаления поврежденных азотистых оснований (BER). Система BER вызывает защиту геномной ДНК от повреждений, вызываемых главным образом алкилирующими агентами, а также эндогенными генотоксическими соединениями, включая внутриклеточные радикалы кислорода и другие реакционноспособные метаболиты, часть из которых уже обсуждалась в начале этой главы. BER начинает функционировать с отщепления ошибочно включенных или модифицированных оснований от дезоксирибозы под действием ключевого фермента – ДНК-гликозилазы, обладающего способностью отщеплять большое число модифицированных оснований ДНК (рис. 1.57). Кроме этих модифицированных оснований в процессе BER может происходить удаление и других производных, образующихся под действием химических мутагенов. В частности, недавно было показано, что по такому же механизму происходит вырезание этнопуриновых производных оснований, образующихся под действием винилхлорида, а также С8-аддуктов аминифлуорена с остатками гуанина. Разные ДНК-гликозилазы благодаря их различной субстратной специфичности осуществляют удаление конкретных модифицированных оснований (табл. 1.20).

Таблица I.20

ДНК-гликозилазы и эндонуклеазы клеток микроорганизмов и человека, участвующие в BER

Фермент	Источник	Ген	Субстрат (см. рис. I.57)
Урацил-ДНК-гликозилаза	<i>E. coli</i>	<i>ung</i>	а
	<i>S. cerevisiae</i>	<i>UNG</i>	»
	Человек	<i>UDG</i>	»
3-Метиладенин-ДНК-гликозилаза	<i>E. coli</i>	<i>tag</i>	к
	»	<i>alkA</i>	з, к–м, (б, и)
	<i>S. cerevisiae</i>	<i>MAG</i>	б, к, л
	Человек	<i>MPG</i>	к, (д)
Фару/8-оксогуанин-ДНК-гликозилаза (фару – формамидопиримидин)	<i>E. coli</i>	<i>fpg/mutM</i>	в–д
	<i>S. cerevisiae</i>	?	г и/или д
	Человек	?	»
Эндонуклеаза III/тимингликоль-ДНК-гликозилаза	<i>E. coli</i>	<i>nth</i>	в, е, ж
Эндонуклеаза VIII	<i>E. coli</i>	<i>nei</i>	»
А-Г-ДНК-гликозилаза	»	<i>mutY</i>	Аденин/в
	Человек	?	»
Г-Т-ДНК-гликозилаза	»	?	Г-Т, (U-G)
УФ-эндонуклеаза	T4	?	Пиримидиновые димеры
	<i>M. luteus</i>	?	То же
	Человек	?	з
Гидроксиметилурацил-ДНК-гликозилаза	Человек	?	з
Формилурацил-ДНК-гликозилаза	»	?	Ж

Примечание. В скобках приведены предположительные субстраты.

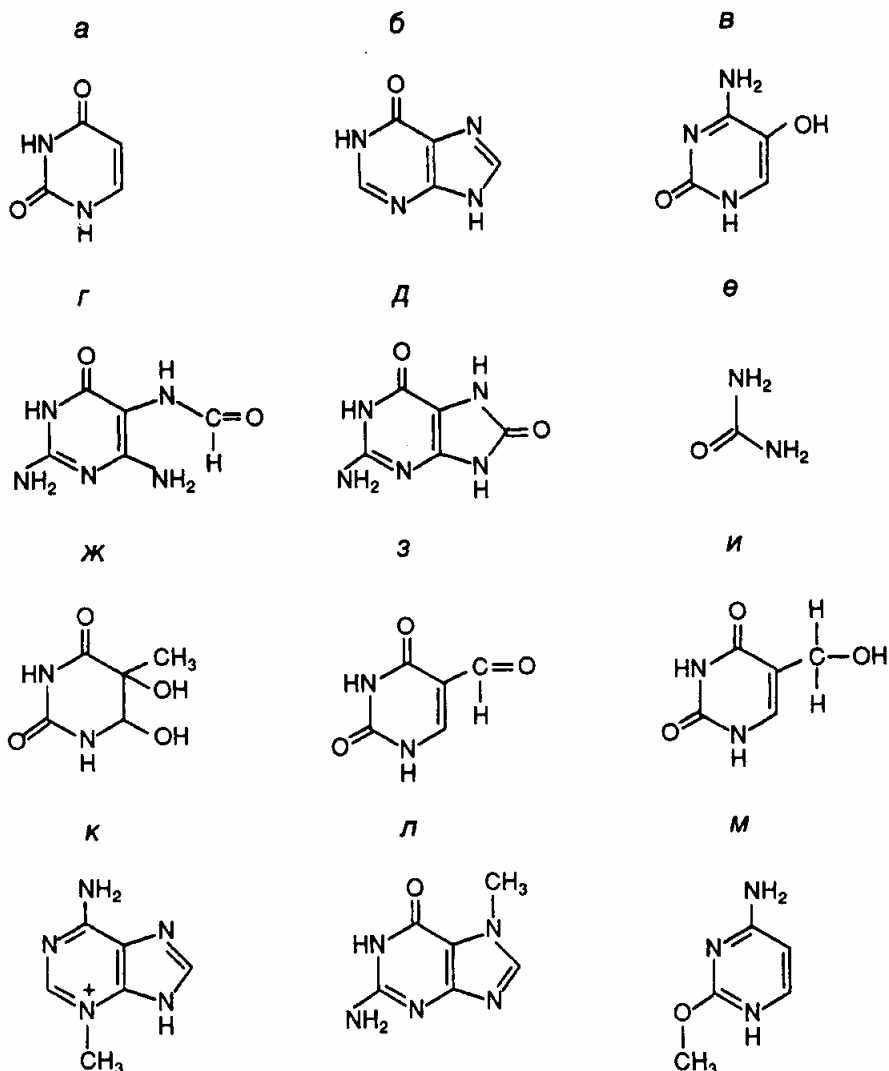


Рис. 1.57. Модифицированные азотистые основания ДНК, удаляемые ДНК-гликозилазами при функционировании BER

а – урацил; б – гипоксантин; в – 5-гидроксицитозин; г – 2,5-диамино-4-формамидопиримидин; д – 7,8-дигидро-8-оксогуанин; е – мочевиная; ж – тимингликоль; з – 5-формилурацил; и – 5-гидроксиметилурацил; к – 3-метиладенин; л – 7-метилгуанин; м – 2-метилцитозин

AP-дезоксирибоза (apurinic/apyrimidinic deoxyribose), образовавшаяся в результате удаления модифицированного азотистого основания апуринового/апиримининового (AP-) сайта, далее вырезается с помощью AP-лиазы, которая освобождает ее 3'-конец, и AP-эндонуклеазы, гидролизующей ее 5'-концевую фосфодиэфирную связь в AP-сайте (см. рис. 1.58).

Однонуклеотидная брешь затем заполняется с помощью ДНК-полимеразы, и фосфодиэфирная связь восстанавливается в реакции лигирования. У *E. coli* репаративный синтез ДНК выполняет ДНК-полимераза I, у дрожжей – ДНК-полимераза δ . Из трех ДНК-лигаз, которыми обладают клетки животных, в BER, по-видимому, участвует ДНК-лигаза III.

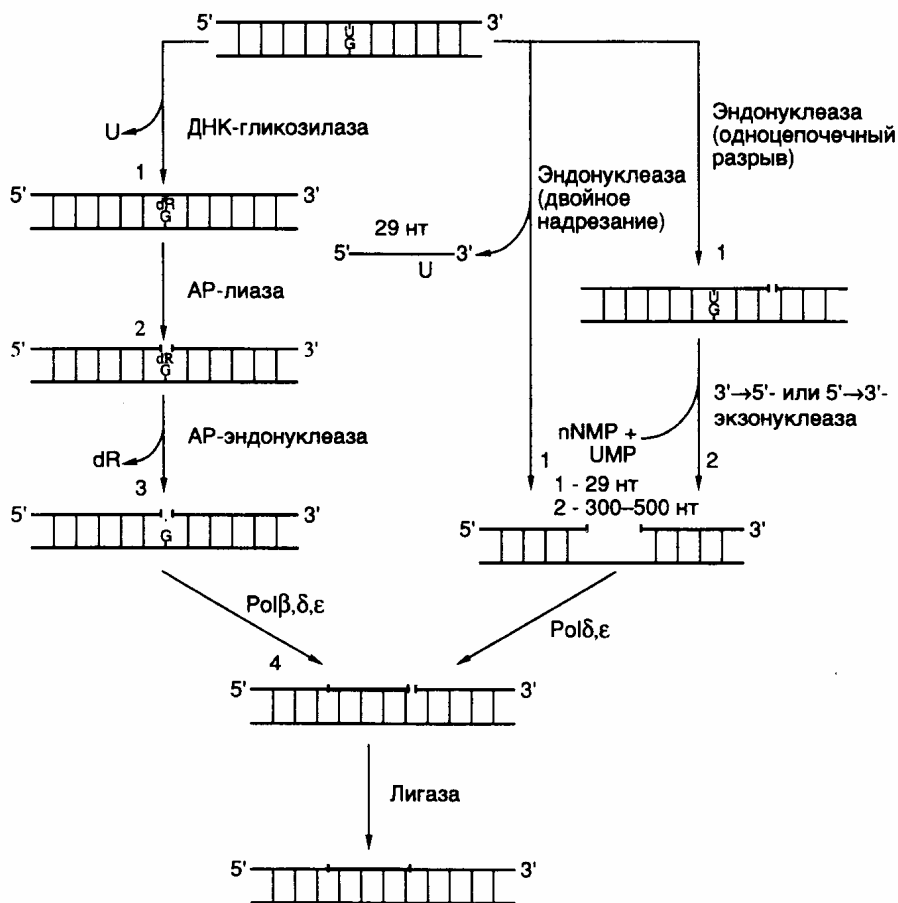


Рис. 1.58. Основные пути и этапы эксцизионной репарации у животных

Цифрами обозначены последовательные этапы функционирования BER и NER

В последнее время начаты исследования механизмов сопряжения BER с другими генетическими процессами, протекающими внутри клеток: транскрипцией, репликацией ДНК и регуляцией клеточного цикла. Для соматических клеток менее опасно иметь повреждения ДНК, связанные с появлением некодирующих (AP-) участков, чем ошибочно кодирующих

оснований, поскольку последние приводят к образованию мутаций, тогда как первые допускают осуществление полноценной пострепликативной репарации повреждений. ДНК-гликозилазы, участвующие в BER, способны переводить сайты, содержащие модифицированные основания (например урацил), в некодирующие сегменты одной из цепей ДНК. Урацилгликозилазы, ассоциированные с белковыми комплексами репликативных вилок, действуют очень эффективно на одноцепочечные ДНК, и их активность регулируется во время клеточного цикла.

Эксцизионная репарация ДНК путем удаления нуклеотидов (NER).

Если в системе BER происходит удаление отдельных поврежденных азотистых оснований ДНК путем разрыва соответствующих N-гликозидных связей между азотистыми основаниями и остатками дезоксирибозы, то в системе NER поврежденные азотистые основания вырезаются в составе олигонуклеотидов. NER может осуществляться двумя путями. В первом случае происходит гидролиз фосфодиэфирной связи по 3'- или 5'-концу на некотором расстоянии от ошибочно спаренного (поврежденного) нуклеотида, который далее целиком удаляется под действием 5'→3'- (или 3'→5'-) экзонуклеазы, гидролизующей цепь ДНК нуклеотид за нуклеотидом в соответствующем направлении от первоначального одноцепочечного разрыва в репарируемой ДНК. Образующаяся брешь далее заполняется ДНК-полимеразой. Такой механизм репарации реализуется у *E. coli* и человека для вырезания неповрежденных (немодифицированных) ошибочно спаренных нуклеотидов. Механизм последовательного эндо- и экзонуклеазного расщепления ДНК не используется для удаления поврежденных (измененных) нуклеотидов. Это связано, по-видимому, с тем, что такие нуклеотиды (например возникшие в результате образования аддуктов с мутагенами) часто являются ингибиторами экзонуклеаз.

Одним из решений данной проблемы представляется использование ферментной системы, которая вносила бы одноцепочечные разрывы по обе стороны от поврежденного нуклеотида на некотором расстоянии от него с последующим удалением одноцепочечного фрагмента ДНК, содержащего измененный нуклеотид. Действительно, такой второй механизм эксцизионной репарации функционирует у всех исследованных видов живых организмов и будет рассмотрен ниже более подробно.

В универсальном механизме эксцизионной репарации как прокариоты, так и эукариоты гидролизуют 3–5-ю фосфодиэфирную связь с 3'-конца от повреждения (см. рис. 1.56). При этом прокариоты гидролизуют также 8-ю связь от 5'-конца измененного нуклеотида, тогда как у эукариотических организмов происходит одноцепочечный разрыв на расстоянии 21–25 нуклеотидов от повреждения со стороны его 5'-конца. Таким образом, прокариоты удаляют измененный нуклеотид в составе 12–13-членных олигомеров, тогда как эукариоты – в составе одноцепочечных фрагментов ДНК длиной в 27–29 нуклеотидов. Ферментная система, вносящая такие двойные одноцепочечные разрывы, получила название *эксцизионной нуклеазы (эксцинуклеазы)*. Образующаяся в молекуле репарируемой ДНК одноцепочечная брешь далее заполняется с помощью ДНК-полимеразы, а фосфодиэфирная связь в остающемся одноцепочечном разрыве восстанавливается ДНК-лигазой.

Генетика NER. Гены NER *E. coli* *uvrA*, *uvrB* и *uvrC* не обнаруживают гомологии с соответствующими генами человека. В отличие от них гены NER дрожжей и человека высокоомологичны, и энзимология эксцизионной репарации в этих двух системах также обладает большим сходством. По крайней мере, три заболевания у человека вызываются генетическими нарушениями системы эксцизионной репарации: пигментная ксеродерма, синдром Кокейна и трихотиодистрофия.

Кожа больных пигментной ксеродермой обладает повышенной чувствительностью к дневному свету, что проявляется в виде фотодерматозов, включая рак кожи. В ряде случаев отмечены аномалии нервной системы, причиной которых являются мутации в одном из семи генов: *XPA*, *XPB*, ...*XPG*. Однако описаны больные с классическими симптомами пигментной ксеродермы, но с ненарушенной системой NER. Для клеток этих больных характерны изменения в так называемой пострепликативной репарации. Больным с синдромом Кокейна присущи нарушения роста, умственная отсталость, катаракты, повышенная чувствительность к свету с сопутствующими дерматозами. Обнаружены мутации в двух группах генов, приводящие к этому заболеванию. У больных с мутантными генами *CS-A* или *CS-B* клетки способны нормально репарировать УФ-повреждения ДНК. У другой группы больных обнаружены мутации в генах *XPB*, *XPD* или *XPG*. У больных трихотиодистрофией со смешанными симптомами выявлены мутации в генах

XPB или *XPD*. Классические симптомы этого заболевания, по-видимому, являются следствием мутации в гене транскрипционного фактора TFIIH.

Получение мутантов с измененной NER у грызунов позволило разбить такие гены на 11 групп комплементации, большинство из которых соответствует группам комплементации XP и CS человека. Часть соответствующих генов человека удалось клонировать, используя их способность исправлять (комплементировать) генетические дефекты в культивируемых мутантных клетках грызунов. Эти гены получили название *кросс-комплементирующих генов* эксцизионной репарации (ERCC – excision repair cross complementing). Среди них гены *XPE* и *ERCC6–ERCC11* не требовались для прохождения основных реакций эксцизионной репарации, и их функция неизвестна.

Структура и функции белков NER. В табл. I.21 суммированы некоторые свойства белков животных, участвующих в NER. Большинство таких белков существует *in vivo* в виде комплексов, поэтому необходимо иметь в виду, что ферментативные активности, обнаруживаемые у отдельных

Таблица I.21

Белки животных, участвующие в NER

Белковая система	Белки системы	Ферментативная активность	Функция в репарации
XPA	XPA (p31)	Связывание ДНК	Распознавание повреждения
RPA	p70 p34 p11	То же	То же
TFIIH	XPB/ERCC3 (p89)	ДНК-зависимая АТРаза	Образование
	XPB/ERCC3 (p89)	Локальное расплетание ДНК	преинициального комплекса
	p62	Фактор транскрипции	
	p44	Киназа, активирующая Cdk	Сопряжение транскрипции и репарации
	Cdk7 (p41) CycH (p38)		

	p34		
XPC	XPC (p125)/ HHR23B (p112)	Связывание ДНК	?
XPF	XPF/ERCC4 (p112) ERCC1 (p33)	Эндонуклеаза	5'-Концевое надреза- ние ДНК
XPG	XPG/ERCC5 (p135)	»	3'-Концевое надреза- ние ДНК

белков в очищенном состоянии, могут не иметь прямого отношения к их функциям в системе NER.

XPA – белок с молекулярной массой 31 кДа, обладает доменом типа "цинковые пальцы", участвует в распознавании поврежденного участка ДНК. Он также взаимодействует с другими компонентами системы и может функционировать в качестве фактора нуклеации для экзонуклеазы. XPA взаимодействует своим N-концевым доменом с гетеродимером ERCC1–XPF, а С-концевым доменом – с TFIIH. Кроме того, белок RPA (HSSB) образует комплекс с XPA и усиливает его специфичность в отношении поврежденной ДНК.

RPA (HSSB) – тример, состоящий из белковых субъединиц p70, p34 и p11, необходим для репликации ДНК и репаративного синтеза, а также для прохождения этапа двойного надреза ДНК во время эксцизионной репарации. Он обладает умеренным сродством к поврежденной ДНК.

TFIIH – олигомерный комплекс, в состав которого входят белки p89, p80, p62, p44, p41, p38 и p34. Этот белковый комплекс первоначально был открыт как один из семи основных факторов транскрипции, необходимых для эффективного функционирования РНК-полимеразы II. Случайно было установлено, что его субъединица p89 идентична белку репаративного комплекса XPB, а также обнаружено отсутствие функциональной комплементации между бесклеточными экстрактами клеток с мутантными белками XPB и XPD, определяемой по восстановлению репарирующей активности в смешанных экстрактах. Все это привело к пониманию того, что весь комплекс TFIIH представляет собой фактор репаративной системы. Белки XPB и XPD являются ДНК-зависимыми АТРазами, обладают так называемыми хеликазными доменами и могут (как и сам фактор TFIIH) вызывать

диссоциацию гибридов, образованных между короткими фрагментами ДНК и одноцепочечной ДНК.

XPC – белок с молекулярной массой ~125 кДа, существует в виде гетеродимера в комплексе с белком p58, который является гомологом белка Rad23 дрожжей (HHR23B). XPC слабо связывается с TFIIH и очень прочно – с одноцепочечной ДНК.

ERCC1/XPF – чрезвычайно прочный белковый комплекс, с которым взаимодействует белок XPA. Он обладает эндонуклеазной активностью, специфичной в отношении одноцепочечной ДНК.

XPG – белковый комплекс, обладающий эндонуклеазной активностью, специфичной в отношении одноцепочечной ДНК; вовлекается в эксцизионный комплекс посредством взаимодействия с TFIIH и RPA.

Механизм NER. Процесс NER условно можно разделить на четыре этапа: а) распознавание поврежденного участка ДНК; б) двойное надрезание (инцизия) цепи ДНК по обеим сторонам поврежденного участка и его удаление (эксцизия); в) заполнение бреши в процессе репаративного синтеза; г) лигирование оставшегося одноцепочечного разрыва ДНК. Феномен NER, как и многие другие генетические явления, имеющие общебиологическое значение, впервые обнаружен у *E. coli*. Было установлено, что мутантные УФ-чувствительные клетки *E. coli* не могут удалять из ДНК тиминовые димеры, возникающие в ответ на действие УФ-света. Вскоре стало ясно, что система эксцизионной репарации не является специфичной в отношении только тиминовых димеров, но способна распознавать и удалять любые повреждения ДНК, возникающие в результате ковалентных модификаций составляющих ее мономеров. Для понимания механизмов узнавания системой эксцизионной репарации поврежденных участков ДНК необходимо ответить, по крайней мере, на три важных вопроса: 1) распознает ли система только поврежденные (модифицированные) основания в ДНК; 2) как система осуществляет выбор цепи ДНК для репарации; 3) каковы молекулярные механизмы распознавания поврежденных участков?

Оказалось, что поврежденные (модифицированные) основания – не единственный субстрат для этой ферментной системы. NER человека распознает и удаляет одиночные ошибочно спаренные нуклеотиды, а также петли длиной в 1–3 нуклеотида. Однако в отличие от истинной репаративной

системы, удаляющей неправильно спаренные основания, NER не может идентифицировать, нуклеотид какой цепи ДНК оказывается правильным. В результате происходит вырезание неспаренных нуклеотидов из любой цепи случайным образом. В отличие от только что рассмотренной ситуации NER человека способна различать цепи ДНК в случае распознавания поврежденных нуклеотидов. В частности, показано, что при наличии в ДНК димеров тимина циклобутанового типа вырезание нуклеотидов происходит исключительно из поврежденной цепи. Механизм такого распознавания в настоящее время неизвестен. К сожалению, остается непонятным и молекулярный механизм узнавания самих поврежденных оснований. Следует заметить, что система способна распознавать повреждения как сильно, так и слабо деформирующие вторичную структуру ДНК. При этом не обнаружена линейная зависимость между коэффициентом специфичности нуклеазы (k_{cat}/k_m) и уровнем деформации двойной спирали ДНК. Показано, что в процессе распознавания участвуют белковые комплексы ХРА/РРА, которые преимущественно связываются с поврежденной ДНК, и TFIIH, обладающий АТР-зависимой ДНК-расплетающей активностью. Последний взаимодействует с поврежденным участком ДНК и по аналогии с соответствующим механизмом у *E. coli* локально раскручивает ДНК, создавая основной преинцизионный комплекс с поврежденной ДНК.

Недавно было установлено, что три фермента репарации, обладающие узкой субстратной специфичностью: ДНК-фотолиаза (удаление пиримидиновых димеров), урацилгликозилаза (удаление урацила из ДНК) и экзонуклеаза III (гидролиз ДНК в AP-сайтах), втягивают поврежденный участок из двойной спирали в полость фермента, что приводит кофактор или аминокислотные остатки активного центра этих ферментов в непосредственный контакт с расщепляемыми связями ДНК. Не исключено, что система эксцизунклеазы действует таким же образом.

Основные этапы функционирования NER, следующие за распознаванием поврежденного участка ДНК, представлены на рис. 1.59. После того как комплекс ХРА–РРА связывается с измененным участком ДНК, ХРА взаимодействует с комплексом TFIIH, который создает преинцизионный комплекс, что сопряжено с гидролизом АТР. АТР-зависимое расплетание ДНК комплексом TFIIH подготавливает ее к взаимодействию с двумя ХР-белками,

обладающими нуклеазной активностью. XPG связывается с TFIIH и вносит одноцепочечный разрыв с 3'-конца повреждения. Аналогично комплекс ERCC1–XPF взаимодействует с XPA в составе преинцизионного комплекса и способствует одноцепочечному разрыву с 5'-конца повреждения. Образование обоих разрывов является АТФ-зависимым, и их расположение на ДНК высокоспецифично. Как правило, происходят разрывы 5-й и 24-й фосфодиэфирных связей соответственно от 3'- и 5'-концов поврежденных участков. Однако расположение точек разрывов может варьировать (см. выше). Таким образом, в результате подобных одноцепочечных надрезов ДНК может освобождаться фрагмент длиной 24–32 нуклеотида с преобладанием фрагментов длиной 27–29 нуклеотидов. На расположение сайтов одноцепочечных разрывов оказывают влияние как характер повреждения, так и последовательности нуклеотидов (контекст), окружающих поврежденный участок. Ту же самую картину инцизии обнаруживают *in vivo* в ооцитах *Xenopus* и у *Schizosaccharomyces pombe*. На этом основании делают вывод об универсальном механизме эксцизионной репарации у эукариот.

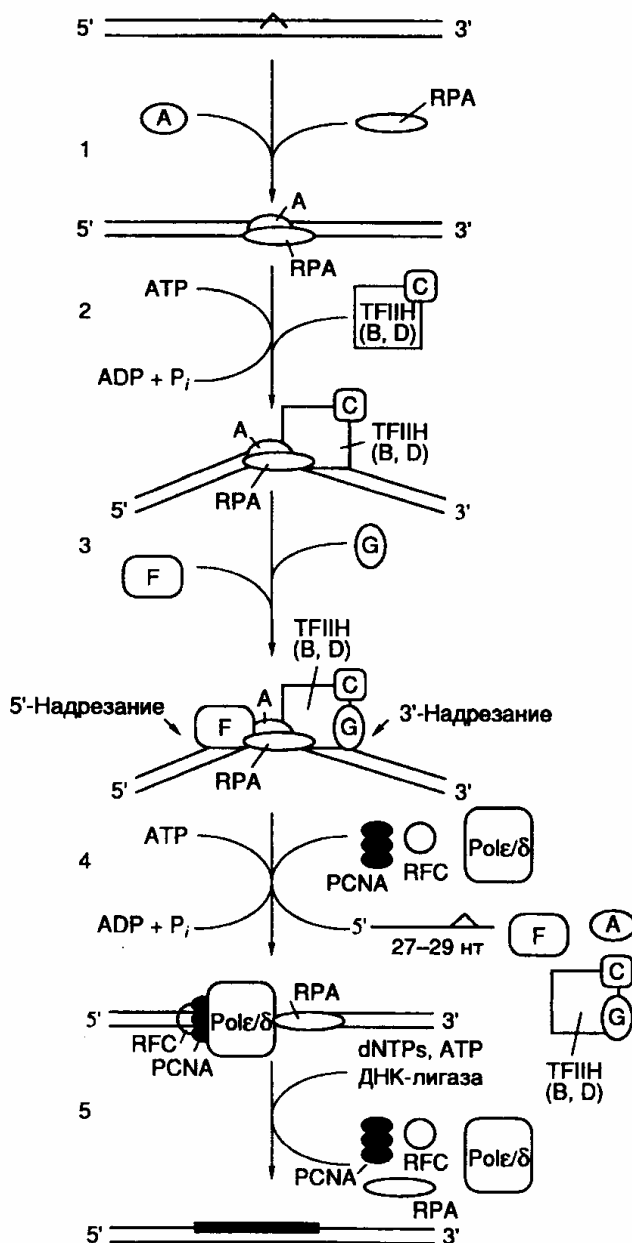


Рис. 1.59. Модель эксцизионной репарации (NER) у млекопитающих
 Обозначены белок-белковые и белково-нуклеиновые взаимодействия, возникающие при функционировании NER. А–F – продукты генов *XPA–XPF*

Репаративный синтез ДНК у человека является PCNA-зависимым (см. раздел 4.1.3), т.е. может осуществляться с участием ДНК-полимераз Pol δ и Pol ϵ . PCNA связывается с системой праймер–матрица под действием фактора

репликации RFC, откуда следует, что последний также участвует в репаративном синтезе ДНК. В опытах с бесклеточными системами моноклональные антитела к Pol δ специфически подавляют репаративный синтез. Однако оказалось, что в тех же высокоочищенных бесклеточных системах вместо Pol δ с аналогичным эффектом могут быть использованы Pol ϵ и даже фрагмент Кленова ДНК-полимеразы I *E. coli*. Это означает, что реконструированные из очищенных компонентов бесклеточные системы лишь в ограниченной степени имитируют биохимические процессы, происходящие в живых клетках. В настоящее время считается, что обе ДНК-полимеразы – Pol δ и Pol ϵ участвуют в репаративном синтезе ДНК у человека.

Сопряжение *NER* с транскрипцией. Транскрибируемые последовательности нуклеотидов ДНК, особенно в матричной цепи, репарируются с большей скоростью, чем нетранскрибируемые последовательности. Интересно, что в клетках больных с синдромом Кокайна не наблюдается такой асимметрии в репарации.

В клетках *E. coli* белковый фактор, кодируемый геном *mfd* и сопрягающий транскрипцию и репарацию, замещает остановившиеся перед повреждением молекулы РНК-полимеразы, что приводит к диссоциации транскрипционного комплекса. При этом он одновременно привлекает экзонуклеазный репаративный комплекс к поврежденному участку ДНК. В клетках животных ген *CSB* кодирует белок с молекулярной массой 160 кДа, который содержит так называемый хеликазный домен (мотив) и, возможно, выполняет те же функции, что и белок Mfd у *E. coli*. На основе поведения клеток с мутантными генами белков CSA и CSB разработана простая модель механизма, с помощью которого обеспечивается асимметричная репарация цепей ДНК. В соответствии с этой моделью РНК-полимераза II, остановившаяся в процессе транскрипции перед поврежденным участком ДНК, распознается комплексом CSA–CSB и перемещается в сторону от повреждения без разрушения четвертичной структуры транскрипционного комплекса. Одновременно комплекс CSA–CSB привлекает компоненты репаративной системы XPA и TFIIH к месту повреждения ДНК и помогает сборке эксцизунклеазного комплекса. Нуклеотиды поврежденной цепи вырезаются, и брешь репарируется. После этого РНК-полимераза в составе транскрипционного комплекса продолжает

транскрипцию.

Регуляция NER. Для клеток животных не характерен так называемый SOS-ответ, свойственный клеткам *E. coli* и представляющий собой суммарную реакцию бактериальной клетки на повреждение ДНК различными агентами, что проявляется, в частности, в усилении транскрипции генов NER. Точно так же посттрансляционные модификации белков репарации, происходящие в ответ на повреждение ДНК, не влияют на активность эксцизунуклеазы человека. Было обнаружено, что повреждения ДНК стабилизируют белок p53 – белок-супрессор опухолевого роста, который является регулятором транскрипции. Имеются данные о том, что белок p53 может взаимодействовать с белками XPB и RPA, необходимыми для NER. Однако клетки с инактивированными генами p53 (p53(-/-)), как и клетки дикого типа, эффективно удаляют из поврежденной ДНК два основных фотопродукта, возникающих под действием УФ-света, и обладают такой же устойчивостью к УФ. Поэтому в настоящее время считается, что белок p53 не оказывает прямого влияния на NER. Белки Cdk7 и циклин H, которые образуют Cdk-активирующую киназу, входят в состав комплекса TFIIH, что позволяет предполагать наличие связи репарации ДНК с фазами клеточного цикла.

5.2.3. Гомологичная рекомбинация в репарации ДНК

Давно известно, что быстро делящиеся бактериальные клетки, содержащие несколько репликонов, образованных недореплицированными хромосомами (см. введение к разделу 4.2), более устойчивы к действию ионизирующей радиации, которая индуцирует двухцепочечные разрывы ДНК, чем клетки с небольшим числом репликонов, находящиеся в стационарной фазе. Гаплоидные клетки дрожжей в фазе G_1 перед началом синтеза ДНК чрезвычайно чувствительны к действию ионизирующей радиации, тогда как те же клетки в фазе G_2 перед митозом так же устойчивы к ионизирующему излучению, как и диплоидные клетки. Эти факты указывают на то, что для эффективного исправления повреждений, вызываемых ионизирующей радиацией, необходимо одновременное присутствие в клетке двух гомологичных молекул ДНК.

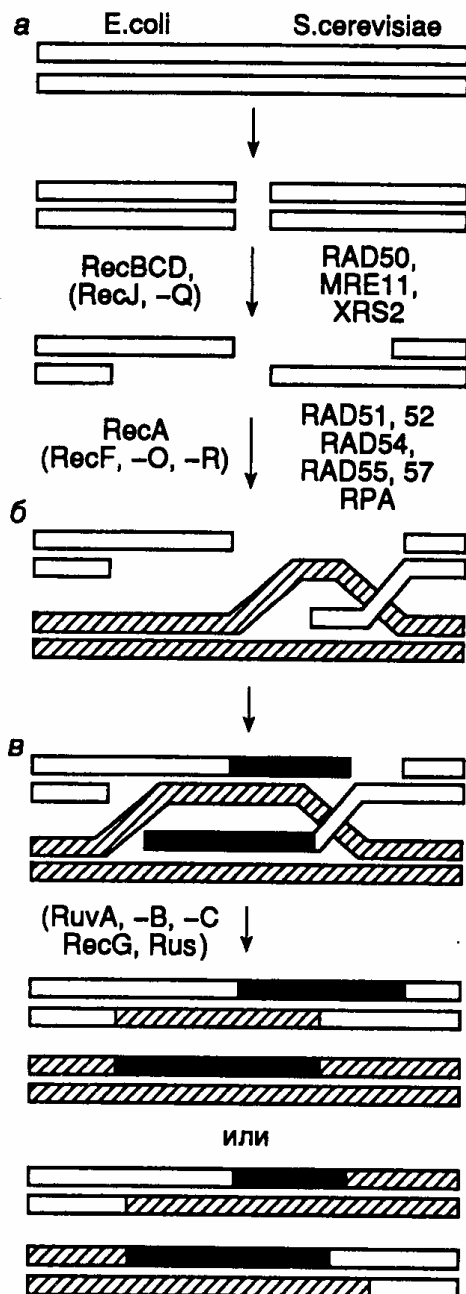


Рис. 1.60. Схема репарации двухцепочечных разрывов ДНК с участием гомологичной рекомбинации

Обозначены продукты генов *E. coli* и *S. cerevisiae*, необходимые для прохождения соответствующих этапов репарации повреждений ДНК. Разделены пресинаптическая (а), синаптическая (б) и постсинаптическая (в) фазы репарации

Существует несколько моделей, объясняющих механизм репарации повреждений ДНК с участием системы гомологичной рекомбинации. Схема одного из таких механизмов репарации двухцепочечных разрывов ДНК у *E. coli* изображена на рис. 1.60. В соответствии с этой моделью процесс репарации условно разделяют на три этапа. В первой, пресинаптической, фазе репарации происходит внесение двухцепочечного разрыва в ДНК или, при его наличии, сразу осуществляется нуклеазное расщепление концов разрыва. В создании одноцепочечных 3'-ОН-выступающих концов ДНК в месте разрыва принимает участие белок RecBCD, который обладает как хеликазной, так и экзонуклеазной активностями. RecBCD расплетает двухцепочечную молекулу ДНК в месте разрыва и гидролизует одну из цепей в направлении 5'→3', оставляя выступающий одноцепочечный участок. Во второй фазе наблюдается синапсис гомологичных участков двух молекул ДНК с вхождением комплементарного одноцепочечного участка в ДНК-дуплекс и последующим репаративным синтезом ДНК. Поиск гомологичных участков и обмен цепями, необходимые для рекомбинации, происходят с участием белка RecA. В третьей, постсинаптической, фазе репарации образовавшиеся структуры Холидея разделяются с помощью белков RuvA, -B и -C, RecG, а также белков SOS-системы репарации (RecN, UvrD, RecF и RecJ). Похожие механизмы используются клетками для рекомбинационной репарации одноцепочечных брешей, остающихся в молекулах ДНК из-за блокировки репликативного синтеза ДНК модифицированными нуклеотидами.

Многие продукты генов *E. coli* и дрожжей, участвующие в рекомбинационной репарации повреждений ДНК, имеют гомологи у животных и человека. Отличительной особенностью эукариотической рекомбинации и репарации является вхождение соответствующих белков в многочисленные белковые комплексы, в частности транскрипсосомы и реплисомы, что указывает на их важную роль в матричном биосинтезе нуклеиновых кислот эукариотических клеток.

5.2.4. Репарация ошибочно спаренных нуклеотидов

Система, осуществляющая репарацию ошибочно спаренных нуклеотидов (mismatch repair), выполняет в клетке несколько важных функций. Прежде всего

она исправляет ошибки репликации ДНК, меняя ошибочно включенные нуклеотиды. Кроме того, при участии этой системы происходит процессинг промежуточных продуктов рекомбинации, приводящий к образованию новых сочетаний генетических маркеров. Ферменты данной системы обеспечивают рекомбинацию между дивергировавшими последовательностями гомологичных ДНК, а также задержку клеточного цикла в ответ на повреждения ДНК. Система репарации ошибочно спаренных нуклеотидов у *E. coli*, использующая белки MutHLS, распознает и репарирует все некомплементарные пары оснований за исключением С–С. Кроме того, эта система репарирует небольшие вставки в одну из цепей ДНК, образующиеся в результате ошибок репликации, длина которых не превышает четырех нуклеотидов.

Обычно у *E. coli* ДНК метилирована Dam-метилазой по сайтам GATC. Однако после завершения репликации дочерняя цепь ДНК некоторое время остается неметилованной. Система MutHLS избирательно репарирует дочернюю цепь ДНК, тем самым значительно повышая точность репликации. Эта система может быть реконструирована *in vitro* с использованием ДНК с одной метилированной цепью в качестве субстрата, к которой добавляются очищенные белки MutH, MutL, MutS, UvrD (хеликаза II), холофермент ДНК-полимеразы III, ДНК-лигаза, белок SSB, а также одна из экзонуклеаз: ExoI, ExoVII или RecJ. Процесс репарации инициируется внесением одноцепочечного разрыва в неметилованную цепь вблизи частично метилированного сайта GATC с последующим гидролизом цепи ДНК и заполнением образующейся одноцепочечной брешы. При этом белок MutS связывается с ошибочно спаренными нуклеотидами. У белка MutL не обнаружено ферментативной активности, хотя он взаимодействует с MutS и необходим для активации MutH – эндонуклеазы, осуществляющей одноцепочечный разрыв ДНК. Таким образом, комплекс MutS–MutL, собранный на участке ДНК с ошибочно спаренным нуклеотидом, стимулирует эндонуклеазную (никазную) активность MutH. Бесклеточная система не требует присутствия MutH при наличии в ДНК-субстрате одноцепочечного разрыва. MutHLS-система репарации может использовать частично метилированные последовательности GATC, расположенные выше и ниже поврежденного участка ДНК. При этом в вырезании ошибочно включенного нуклеотида помимо хеликазы II принимает участие одна из экзонуклеаз: ExoI (3'-экзо), ExoVII (3'- и 5'-экзо) или RecJ (5'-

экзо) в зависимости от расположения GATC-сайта по отношению к корректируемому нуклеотиду. Вслед за вырезанием нуклеотида образовавшаяся одноцепочечная брешь заполняется холоферментом ДНК-полимеразы III в присутствии SSB-белка и ДНК-лигазы.

Следует подчеркнуть, что использование белка MutH и Dam-метилазы для распознавания дочерней цепи реплицировавшейся ДНК является уникальным свойством грамотрицательных бактерий. У грамположительных бактерий не происходит метилирование цепей ДНК в целях маркировки. Если сайты GATC полностью метилированы, MutHLS-система репарации *E. coli* изменяет ошибочно спаренные нуклеотиды в обеих цепях ДНК с одинаковой эффективностью.

У *E. coli* существуют, по крайней мере, еще два специфических пути репарации ошибочно спаренных нуклеотидов. Система VSP (very short patch repair pathway) репарирует некомплементарные пары G–T, заменяя их на G–C. Считается, что такие пары образуются в результате дезаминирования 5-метилцитозина в сайтах, где остатки C метилированы Dcm-метилазой. С более низкой эффективностью эта же система заменяет пары G–U на G–C. Другая MutY-зависимая система репарации специфически ликвидирует последствия окислительных повреждений гуанина. Если dGTP окисляется с образованием 8-оксо-dGTP, белок MutT расщепляет последний, предотвращая его включение в ДНК. Если же он все-таки включается напротив остатка C, то Fpg-гликозилаза (MutM) удаляет это модифицированное основание. В том случае, когда 8-оксо-G остается в составе ДНК, в следующем раунде репликации он спаривается с A, и в итоге может произойти трансверсия G–C→T–A. В этом случае белок MutY действует как ДНК-гликозилаза, удаляющая остаток A из некорректной пары, и как AP-лиаза, вносящая одноцепочечный разрыв по соседству с AP-сайтом. Далее следуют процессы, уже рассмотренные выше в связи с функционированием системы репарации BER. Последовательность реакций с участием MutY также репарирует некомплементарные пары A–G и A–C с образованием соответственно пар C–G и G–C.

Репарация ошибочно спаренных оснований у эукариот происходит при участии комплекса белков, подобного системе MutHLS бактерий. Белок GTBP человека представляет собой гомолог бактериального белка MutS, а у дрожжей в соответствующей роли выступает белок Msh6. Распознавание ошибочно

спаренных нуклеотидов у человека осуществляется гетеродимером MSH2–GTBP. Гомологами MutL в клетках *S. cerevisiae* являются белки MLH1 и PMS2, которые также существуют в виде гетеродимерных комплексов. Мутации в генах, кодирующих эти белки у человека, сопровождаются формированием мутаторного фенотипа и развитием наследственного неполипозного рака кишечника (синдром HNPCC – hereditary nonpolyposis colon cancer).

5.2.5. Полимераза поли(ADP-рибозы) в репарации ДНК у эукариот

В отличие от бактерий одним из первых ответов клеток животных на тяжелые повреждения ДНК является массированная полимеризация остатков ADP-рибозы специальным ферментом – *полимеразой поли(ADP-рибозы)* (poly(ADP-ribose polymerase) – PARP). В ядрах клеток млекопитающих PARP присутствует в количестве $\sim 10^6$ копий, и она обнаружена у большинства эукариот за исключением дрожжей. Процесс синтеза поли(ADP-рибозы) предшествует началу репарации повреждений ДНК. Большие затраты энергии на биосинтез этого полимера указывают на его важную, хотя и до конца не понятную, роль в выходе ядер клеток из стрессового состояния, вызванного премутационными повреждениями ДНК.

In vivo поли(ADP-рибоза) характеризуется очень малым временем полужизни. Цепи полимера, синтезированные в ядрах в ответ на мутагенное воздействие, в основном распадаются уже через 1–2 мин после завершения их синтеза. Такой быстрый обмен полимера в ядрах становится возможным благодаря совместному действию двух ферментов – PARP и гликозилазы поли(ADP-рибозы). Экспрессия PARP-кДНК в ядрах дрожжей, у которых отсутствуют оба фермента, летальна для клеток. Это связано, главным образом, с тем, что внутриядерное накопление поли(ADP-рибозы) сопровождается подавлением репликации ДНК и транскрипции.

Ген PARP, картированный на участке 1q41–1q42 хромосомы человека, кодирует полипептид, состоящий из двух функционально различающихся частей: N-концевого ДНК-связывающего и С-концевого каталитического доменов. Первый домен содержит две структуры типа "цинковые пальцы", которые обеспечивают взаимодействие PARP с разрывами в ДНК. С помощью футпринтинга установлено, что PARP связывается преимущественно с

одноцепочечными разрывами ДНК, закрывая своей полипептидной цепью по семь–восемь нуклеотидов по обе стороны от разрыва. При этом PARP индуцирует образование V-образного изгиба ДНК в месте одноцепочечного разрыва. Для синтеза поли(ADP-рибозы) фермент использует NAD в качестве субстрата. Структурные аналоги NAD часто являются ингибиторами PARP, например 3-аминобензамид – один из самых эффективных ингибиторов PARP. Каталитический домен PARP обнаруживает гомологию с различными NAD-связывающими ферментами. Одноцепочечные разрывы, остающиеся между фрагментами Оказаки при синтезе отстающей цепи ДНК, не индуцируют образование поли(ADP-рибозы), вероятно, из-за их экранирования белками реплисомы.

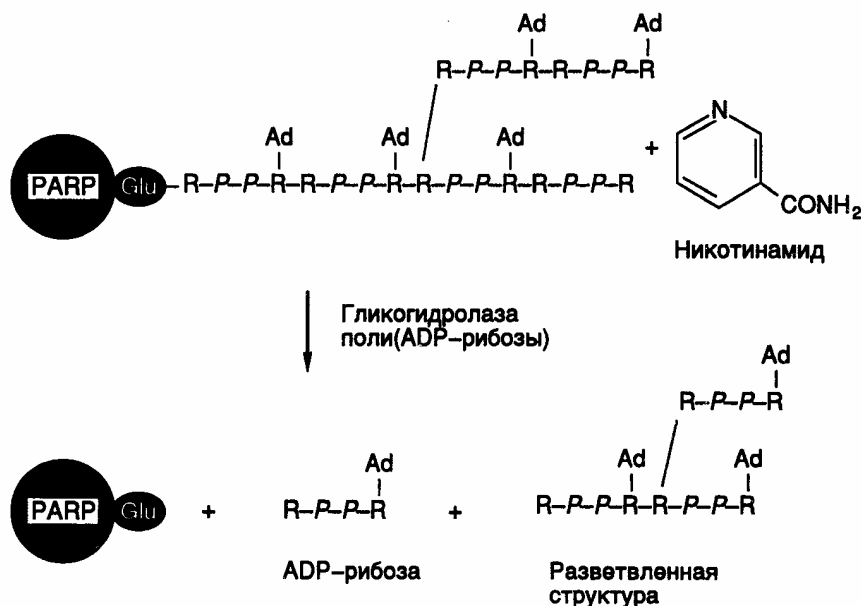


Рис. 1.61. Схематическое изображение молекулы поли(ADP-рибозы), присоединенной к полипептидной цепи полимераза поли(ADP-рибозы) и продуктов ее деградации

Синтез поли(ADP-рибозы) можно рассматривать как один из редких случаев посттрансляционной модификации белков, при которой PARP использует свою собственную полипептидную цепь в качестве субстрата (рис. 1.61). Такая массивная аутомодификация резко изменяет физические свойства фермента. Остатки Glu (25–30), расположенные в полипептидной цепи PARP между двумя вышеупомянутыми доменами, служат точками

инициации синтеза поли(ADP-рибозы). В процессе синтеза происходит разветвление полимера, и длина боковых цепей может достигать нескольких сотен остатков. В настоящее время до конца неизвестно, модифицирует ли молекула PARP сама себя или же это осуществляют другие молекулы PARP. На основании имеющихся кинетических данных наиболее вероятной считается модель, в соответствии с которой молекула PARP, ассоциированная с одноцепочечным разрывом, образует комплекс с другой молекулой и уже вторая молекула PARP служит акцептором полимеризуемой ADP-рибозы.

Гликогидролаза поли(ADP-рибозы) расщепляет цепи полимера с их концов, освобождая мономеры и олигомеры ADP-рибозы. Таким образом, в результате синтеза и деградации поли(ADP-рибозы) в ядрах образуются не только остатки никотинамида и ADP-рибозы, но и более сложные разветвленные продукты, состоящие из трех–четырёх молекул мономера.

В присутствии ингибиторов PARP клетки животных становятся чрезвычайно чувствительными к действию алкилирующих агентов и ионизирующей радиации. Кроме того, в этом случае наблюдается повышенный уровень сестринских хроматидных обменов (СХО). Однако до сих пор нет доказательств прямого участия поли(ADP-рибозы) в репарации ДНК. Сверхэкспрессия рекомбинантного N-концевого домена PARP в клетках животных сопровождается теми же эффектами, что и действие ингибиторов PARP, в частности 3-аминобензамида. В неповрежденных клетках происходят дестабилизация генома и рост числа СХО. Опыты с трансгенными мышами, у которых ген PARP инактивирован в результате генного нокаута (см. раздел 10.3.4), показывают, что PARP-дефицитные мыши, тем не менее, здоровы и фертильны. Следовательно, PARP не играет существенной роли в пролиферации клеток, их дифференцировке и онтогенезе мыши. Клетки таких животных обладают нормальной способностью к эксцизионной репарации ДНК системами BER и NER. PARP-дефицитным мышам свойственны не совсем понятные физиологические дефекты, проявляющиеся в гиперплазии эпидермиса, вызванной повышенной пролиферацией кератиноцитов у старых особей, а также пониженной способностью тимоцитов к пролиферации после γ -облучения. Клетки панкреатических островков у PARP-дефицитных мышей обладают повышенной устойчивостью к цитотоксическому действию NO. Это может указывать на недостаток NAD, вызываемый повышенным синтезом

поли(ADP-рибозы), как одну из причин цитотоксичности NO у нормальных животных.

В настоящее время предложено несколько моделей, объясняющих физиологическую роль PARP в клетках животных. Все они подчеркивают, что PARP не участвует прямо в эксцизионной репарации ДНК, но необходима для быстрой мобилизации ресурсов клеток при исправлении повреждений их генома. Способность ингибиторов PARP вызывать гиперчувствительность клеток к алкилирующим агентам и ионизирующей радиации позволяет рассматривать этот фермент в качестве удобной мишени в химиотерапии опухолей.

Рассмотренные в данной главе механизмы образования мутаций указывают на большое разнообразие путей повреждения генетической информации, заключенной в экспрессирующихся и временно молчащих генах. Эволюционное развитие животного и растительного мира противопоставило мутагенным воздействиям мощное противоядие в виде эффективных систем репарации ДНК. Тем не менее, из-за ошибок систем репарации и повреждения самих этих систем происходит необратимое накопление мутаций, приводящих к нарушениям метаболизма и развитию различных патологических состояний организма. В этой связи еще более эффективным средством защиты генетической информации является блокирование (инактивация) химических мутагенов на подступах к жизненно важным генетическим локусам. Именно такую нагрузку несут рассмотренные выше ферментные системы детоксикации ксенобиотиков. На мой взгляд, одной из функций поли(ADP)-рибозы в клетках, подвергнутых мутагенному воздействию, может быть очистка ядер от мутагенов, ковалентно взаимодействующих с нуклеиновыми кислотами, в том числе свободных радикалов, возникающих под действием ионизирующего излучения. Это соединение может играть роль чистильщика (scavenger) ядер от химических мутагенов и понижать их внутриядерную концентрацию через образование соответствующих аддуктов. Синтез поли(ADP)-рибозы является быстрым адаптивным ответом клеток в ответ на массированное мутагенное воздействие. Однако в естественных условиях существования организмов такие ситуации встречаются редко. В природных условиях, свободных от присутствия антропогенных экологических факторов, скорее имеется слабый мутагенный фон, постоянно окружающий информационные макромолекулы на

протяжении всей жизни организма. В недавно разработанной модели альтруистичной ДНК (Л.И. Патрушев, 1997 г.) избыточные последовательности нуклеотидов эукариотической ДНК рассматриваются в качестве еще одной специфической системы защиты генетической информации, характерной для многоклеточных организмов. Далее будут представлены основные положения этой модели.

5.3. Альтруистичная ДНК

Как следует из вышеизложенного, стабильность генетической информации любого организма обеспечивается двумя различными путями. Прежде всего, системы детоксикации ксенобиотиков и эндогенных мутагенов осуществляют блокирование токсического потенциала таких соединений путем их химических модификаций, обеспечивающих эффективное выведение из клетки и организма. С другой стороны, если генотоксические агенты, прорвав этот барьер, все-таки модифицируют ДНК, начинает функционировать многокомпонентная система репарации поврежденных генов.

Анализ структуры генома эукариот показывает, что для соматических клеток многоклеточного организма остается еще один путь защиты своего генома от мутагенов экзогенного и эндогенного происхождения – разбавление кодирующих последовательностей нуклеотидов некодирующими таким образом, чтобы последние выступали в роли ловушек мутагенов, будь то химические мутагены, непосредственно взаимодействующие с ДНК, или же модифицированные нуклеотиды с измененной специфичностью спаривания, ошибочно включаемые ДНК-полимеразами в ДНК из внутриклеточного пула модифицированных предшественников в процессе репликации. В живом организме, обеспечивающем поддержание внутриклеточного гомеостаза, внутриядерная среда соматических и половых клеток, включающая низкомолекулярные метаболиты, на протяжении больших промежутков времени (соизмеримых со временами клеточного цикла или даже продолжительности жизни самого организма) должна находиться в стационарном состоянии. Можно полагать, что в обычных экологических условиях, к которым организм адаптирован, внутриядерные мутагены, поступающие из цитоплазмы или образующиеся в самом ядре, находятся в стехиометрическом недостатке по отношению к потенциальным мишеням,

способным их акцептировать, в том числе и азотистым основаниям геномной ДНК. Кроме того, в первом приближении можно считать, что расположенные по соседству кодирующие и некодирующие последовательности нуклеотидов ДНК в равной степени доступны действию на них внутриядерных мутагенов. В таких условиях вероятность образования аддуктов внутриядерных мутагенов с кодирующими последовательностями нуклеотидов геномной ДНК будет прямо пропорциональна их суммарной длине (доле) в геномной ДНК или обратно пропорциональна "разбавлению" этих последовательностей некодирующими последовательностями нуклеотидов. Такое разбавление могло бы произойти в результате эволюционных преобразований генома-предшественника путем включения в него некодирующих избыточных последовательностей при участии разных молекулярных механизмов.

Вовлечение систем репарации повреждений ДНК в поддержание генетической стабильности информационных макромолекул клетки является вынужденной мерой, указывающей на то, что действие всех остальных систем защиты не обеспечивает полной сохранности нативного состояния генома. Тем не менее, любая генетическая система будет функционировать надежнее в том случае, если наиболее важные в обеспечении жизнедеятельности генетические локусы будут дополнительно защищены от действия химических мутагенов. Действительно, с точки зрения защиты информации гораздо большего эффекта можно достичь путем создания дополнительных препятствий на пути генотоксических агентов к таким локусам, чем репарацией поврежденных генов, так как в последнем случае не всегда повреждение может быть исправлено и исходная первичная структура ДНК восстановлена. Как следует из дальнейшего изложения, любому, и особенно эукариотическому, геному свойственна дифференциальная защита индивидуальных генетических локусов с помощью очень простого механизма. Необходимость дополнительной защиты генетической информации особенно актуальна для многоклеточных организмов в связи с тем, что у них существует опасность накопления соматических мутаций во время онтогенетического развития, когда создаются гигантские клоны высококооперированных и специализированных соматических клеток.

5.3.1. Парадокс возможности существования многоклеточных организмов

Огромный размер генома многоклеточных организмов с генетической

точки зрения должен создавать для их существования многочисленные и, на первый взгляд, трудноразрешимые препятствия. Проблемы начинаются уже при редупликации гигантских молекул геномной ДНК с помощью ферментных систем, точность функционирования которых не является абсолютной. Кроме того, репликация происходит в присутствии мутагенов экзогенного и эндогенного происхождения. Принято считать, что частота спонтанных мутаций в геноме соматических клеток млекопитающих, растущих в культуре, варьирует от локуса к локусу одного и того же генома и, по разным оценкам, достигает значений от 10^{-8} до 10^{-12} на нуклеотид за клеточную генерацию. Точное определение частоты спонтанных мутаций в геномной ДНК высших эукариот *in vivo* является сложной задачей. По ряду экспериментальных и косвенных данных, в частности, исходя из частоты встречаемости некоторых наследственных заболеваний в популяциях человека (например серповидно-клеточной анемии, вызываемой заменой единственного нуклеотида в геномной ДНК), полагают, что эти показатели, по крайней мере, не ниже значений, полученных *in vitro*.

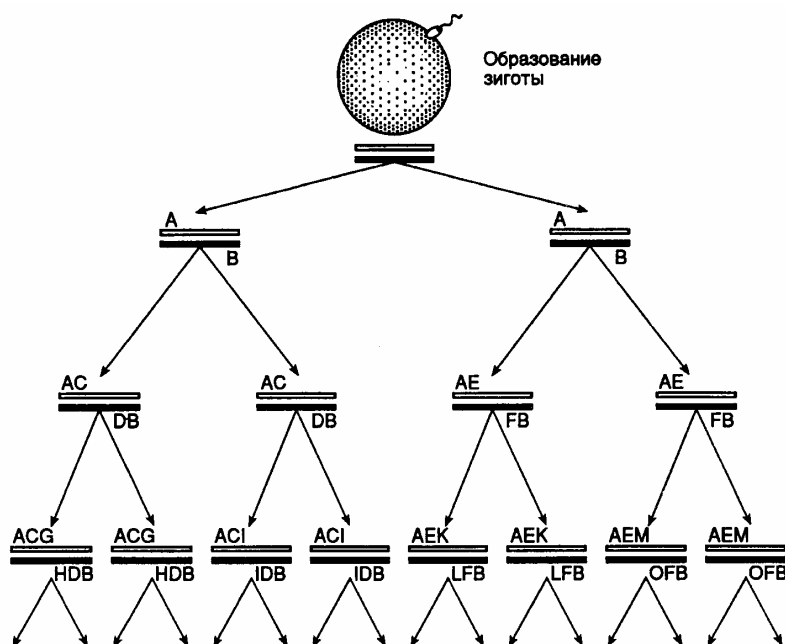


Рис. 1.62. Происхождение генетического груза в геноме соматических клеток многоклеточных организмов

Развитие многоклеточного организма начинается с дробления зиготы, образующейся в процессе оплодотворения яйцеклетки и содержащей

диплоидный набор хромосом (две параллельные линии вверху рисунка).
 А–М – наборы мутаций

Принимая, что суммарная ДНК гаплоидного генома человека насчитывает $\sim 3 \cdot 10^9$ п.о., а частота спонтанных мутаций в среднем составляет 10^{-8} на нуклеотид за генерацию, можно предположить, что, начиная с первого деления оплодотворенной яйцеклетки в процессе онтогенетического развития организма человека, каждое следующее деление должно сопровождаться появлением в их геномной ДНК, по крайней мере, 30 независимых мутаций (рис. 1.62). Организм человека состоит из $\sim 10^{15}$ клеток. Для образования стольких клеток из оплодотворенной яйцеклетки требуется ~ 50 клеточных генераций. Следовательно, гаплоидный геном каждой из соматических клеток человека 50-й генерации должен содержать в разных частях, по крайней мере, 1500 мутаций. Если предположить далее, что набор из 30 мутаций возникает и закрепляется в каждом последующем клеточном делении независимо от мутаций, полученных во время предыдущих делений клеток, то получается, что любая дочерняя соматическая клетка наследует от клетки-предшественницы все имеющиеся в ее геноме мутации и приобретает блок новых 30 мутаций. При этом набор соматических мутаций в потомстве каждой из делящихся клеток одного поколения будет отличаться один от другого, в том числе и в парах аллельных локусов ДНК отдельных клеток, поскольку они редуцируются независимо. При таком развитии событий мутации, возникающие в геноме соматических клеток каждой последующей генерации, случайным образом сканируют шаг за шагом всю реплицирующуюся геномную ДНК, а результаты сканирования не повторяются в каждом новом поколении клеток. Так, блок из 30 спонтанных мутаций, возникающих после гипотетической 50-й генерации соматических клеток в онтогенезе многоклеточного организма, должен присутствовать в $\sim 10^{15} (2^{50})$ вариантах.

Проблема становится еще более очевидной, если иметь в виду, что многие клетки организма, например эпителиальные или стволовые клетки крови, пролиферируют на протяжении всей жизни многоклеточного организма, совершая громадное число клеточных делений. По некоторым оценкам, общее число клеточных циклов, в которых участвуют клетки человека на протяжении его жизни, приближается к 10^{16} . Кроме того, соматические мутации возникают в

многоклеточном организме не только в активно пролиферирующих, но и покоящихся клетках. Частоты возникновения спонтанных мутаций в пролиферирующих и в покоящихся клетках млекопитающих в ряде случаев различаются лишь незначительно.

Парадоксальность эволюционно сложившейся генетической ситуации заключается в том, что если бы большая часть последовательностей нуклеотидов геномной ДНК заключала в себе жизненно важную генетическую информацию, существование многоклеточных организмов было бы невозможно. Их гибель происходила бы из-за неизбежного накопления в делящихся соматических клетках вредных или летальных мутаций, приводящих к обрыву линий дифференцирующихся в онтогенезе соматических клеток. Это особенно относится к мутациям в жизненно важных генах половых хромосом, которые в соматических клетках находятся в гемизиготном состоянии. Альтернативно основные генетические локусы многоклеточных организмов могут находиться под дополнительной защитой от мутационных изменений.²

На основании приведенных выше аргументов можно сделать вывод о том, что большая часть некодирующих избыточных последовательностей нуклеотидов геномной ДНК эукариот не заключает в себе жизненно важной генетической информации и/или у эукариотических клеток имеются механизмы, обеспечивающие дополнительную защиту геномной ДНК от мутаций. Не соглашаясь с основными положениями концепции "эгоистической" и "паразитической" ДНК, можно предположить, что именно такая организация генома эукариотических организмов, в корне отличающаяся от структуры генома прокариот, имеет отношение к разрешению проблемы генетической

³ Нарисованная выше картина, как и рис 1.62, являются упрощенной схемой, в которой, в частности не учитывается рецессивный характер большинства возникающих мутаций. Однако наличие генетического груза в популяциях многоклеточных организмов, а также гемизиготное состояние части генов, ассоциированных с половыми хромосомами, повышают вероятность полного функционального выключения аллелей под действием соматических мутаций, возникающих в онтогенезе диплоидных организмов. Дублирование генетической информации (тотипотентность соматических клеток), позволяющее производить замещение поврежденных клеточных линий в онтогенезе, и другие защитные механизмы, которые будут обсуждаться ниже, снижают остроту проблемы соматических мутаций в онтогенезе многоклеточных организмов, но не снимают ее полностью. Такой проблемы не возникает у прокариот из-за отсутствия у них сомы, а следовательно, и

парадоксальности существования многоклеточных организмов, которую можно назвать *парадоксом М* (Metazoa, Metaphyta). Создается впечатление, что именно избыточная ДНК генома эукариот может иметь отношение к повышению его информационной стабильности до уровня, необходимого для реализации многоклеточности в природе.

5.3.2. Повышение информационной стабильности генома избыточными последовательностями

Анализ структуры генома современных эукариот показывает, что эволюционные преобразования генома-предшественника, приведшие к включению в него избыточных последовательностей нуклеотидов, сопровождались важными генетическими изменениями в отношении стабилизации генетической информации. В частности, многократное превышение содержания избыточных последовательностей нуклеотидов над кодирующими неизбежно должно приводить к соответствующему уменьшению вероятности возникновения мутаций в кодирующих и других функционально значимых частях под действием внутриядерных мутагенов эндогенного и экзогенного происхождения. Поскольку в разных частях интерфазного ядра (микрокомпартаментах, заключающих в себе хромомеры интерфазных хромосом) наблюдается гетерогенность в уровнях конденсации хроматина, индивидуальные локусы могут быть по-разному защищены от мутационных изменений, вызываемых мутагенами. Рассмотрим более подробно влияние избыточных последовательностей ДНК генома эукариот на стабильность их генома.

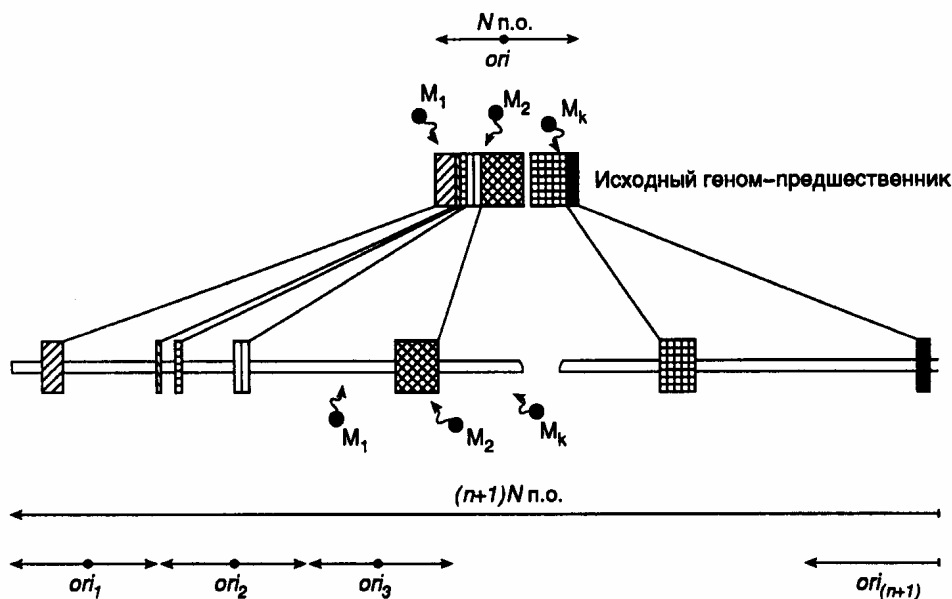


Рис. 1.63. Гипотетическое эволюционное преобразование генома-предшественника путем включения в него некодирующих избыточных последовательностей нуклеотидов

Влияние избыточных последовательностей нуклеотидов на число мутаций, возникающих в результате ошибок репликации в кодирующих последовательностях генома. Предположим, что длина исходного генома, не содержащего избыточных последовательностей нуклеотидов, составляет N п.о. (см. рис. 1.63). При этом в результате ошибок репликации в нем, в среднем, возникает a мутаций независимо одна от другой и случайным образом. Допустим, что в ходе эволюционных преобразований в него включаются избыточные последовательности нуклеотидов, суммарная длина которых составляет nN п.о. и, соответственно, общая длина преобразованного генома становится равной $(n+1)N$ п.о. Поскольку число мутаций, возникающих в результате ошибок репликации, прямо пропорционально длине реплицирующейся ДНК, общее количество мутаций, в среднем, возникающих в преобразованном геноме при участии этого механизма, должно возрасти в $n+1$ раз и составить $a(n+1)$. Вероятность возникновения одной независимой и случайной мутации в некодирующей части генома $P(1)$ будет пропорциональна его длине:

$$P(1) = \frac{nN}{(n+1)N} = \frac{n}{n+1}. \quad (1)$$

В то же время вероятность возникновения в избыточных частях генома всех $a(n+1)$ мутаций будет равна:

$$P[a(n+1)] = \left(\frac{n}{n+1} \right)^{a(n+1)} \quad (2),$$

поскольку вероятность одновременного наступления $a(n+1)$ независимых событий равна произведению вероятностей наступления каждого из них в отдельности. При $a=1$ (т.е. в том случае, если в процессе репликации исходного генома в нем, в среднем, возникала одна мутация) и достаточно больших значениях n это выражение стремится к e^{-1} , т.е. к $\sim 0,36$. Таким образом, в данном случае при $n=100$ (что, приблизительно, соответствует соотношению некодирующих и кодирующих последовательностей нуклеотидов в геноме человека и других млекопитающих) вероятность того, что ни одна из мутаций, возникающих в гипотетическом преобразованном геноме в результате ошибок репликации, не произойдет в его кодирующих частях, будет довольно высокой и составит $\sim 0,37$. Это означает, что, в среднем, каждая третья дочерняя соматическая или половая клетка, возникшая в результате редупликации гипотетического генома с достаточным количеством некодирующих избыточных последовательностей нуклеотидов, будет полностью свободна от мутаций, образующихся по такому механизму в кодирующих частях своего генома.

С увеличением числа мутаций в исходном геноме-предшественнике ($a \gg 1$) вероятность возникновения всех мутаций в некодирующих частях генома быстро уменьшается. Однако поскольку эти мутации будут распределяться между кодирующими и некодирующими частями генома пропорционально длине каждой из этих частей, общее их количество в кодирующих частях генома останется неизменным. Следовательно, эволюционное включение в геном-предшественник большого количества некодирующих последовательностей нуклеотидов не увеличивает число мутаций, возникающих в кодирующих частях генома в результате ошибок репликации. Более того, в ряде случаев такое эволюционное преобразование

генома может заметно стабилизировать его генетическую информацию.

Влияние избыточных последовательностей нуклеотидов на число мутаций, возникающих в кодирующих частях генома под действием мутагенов. Ситуация, связанная с возникновением мутаций в гипотетическом геноме под действием мутагенов экзогенного и эндогенного происхождения, принципиально отличается от только что рассмотренной (см. рис. 1.63). Если предположить, что эволюционное преобразование генома, приведшее к включению в него n некодирующих последовательностей нуклеотидов, не сопровождается увеличением числа внутриядерных мутагенов, то генетические последствия такого преобразования будут гораздо более значительными.

Так же как и в предыдущем случае, вероятность попадания одного мутагена M_k в некодирующую область гипотетического генома равна $\frac{n}{n+1}$. Вероятность же того, что все k мутагенов попадут в некодирующие области нового генома $P(k)$, равна:

$$P(k) = \left(\frac{n}{n+1} \right)^k \quad (3).$$

При больших значениях n и малых k величина $P(k)$ стремится к 1,0, т.е. имеет место событие, близкое к достоверному. Иными словами, чем больше доля некодирующих последовательностей нуклеотидов в геномной ДНК, тем вероятнее, что все внутриядерные мутагены попадут в некодирующие последовательности. В случае гипотетического генома с $n = 100$ вероятность попадания одного мутагена в кодирующую область становится равной 0,01, т.е. весьма малой. При этом общее число мутагенов, которые будут взаимодействовать с кодирующими последовательностями нуклеотидов, уменьшится в 100 раз и будет иметь место 100-кратная защита кодирующих функционально значимых участков эволюционно преобразованного генома от мутаций, вызываемых внутриядерными мутагенами, по сравнению с исходным геномом-предшественником. При $k = n+1$ с ростом n уравнение (3) будет стремиться к e^{-1} , т.е. к $\sim 0,36$, и быстро уменьшаться при дальнейшем увеличении k . Но это относится к вероятности *полной* защиты кодирующих последовательностей нуклеотидов. Относительная же защита, равная доле химических мутагенов из всего их пула, взаимодействующих с кодирующими

последовательностями нуклеотидов, будет обратно пропорциональна общей длине избыточных последовательностей в преобразованном геноме, т.е. обратно пропорциональна n . Относительная защита генома от химических мутагенов может быть особенно актуальной в условиях экологического стресса.

Насколько соответствует действительности предположение о том, что эволюционное преобразование генома-предшественника путем включения в него большого количества избыточных последовательностей нуклеотидов не будет сопровождаться пропорциональным возрастанием количества внутриядерных мутагенов? Очевидно, что такое предположение является упрощением. Увеличение внутриядерного содержания ДНК, например в результате эндорепликации при политении, по-видимому, всегда приводит к пропорциональному увеличению объема соответствующих ядер, следствием чего, казалось бы, должно быть пропорциональное возрастание количества молекул внутриядерных мутагенов. Однако это не совсем так. Данный вывод относится лишь к мутагенам, непосредственно образующимся в ядре, например в результате взаимодействия ионизирующего излучения с веществом ядер. Большая же часть мутагенов, по-видимому, должна поступать в ядра из цитоплазмы путем радиальной диффузии через ядерные мембраны. При этом одним из факторов, ограничивающих попадание мутагенов из цитоплазмы в ядра, является их поверхность, поскольку вероятность контакта мутагенов цитоплазмы с ядром прямо пропорциональна площади его поверхности. При увеличении объема поверхность ядра-шара возрастает пропорционально квадрату его радиуса, тогда как объем – пропорционально кубу радиуса ядра. Следовательно, объем ядер будет увеличиваться быстрее площади их поверхности и количество молекул внутриядерных химических мутагенов, приходящихся на нуклеотид ядерной ДНК, должен уменьшаться при возрастании ядерного объема за счет увеличения содержания ядерной ДНК. Такое эволюционное преобразование генома в целом будет сопровождаться повышением его информационной стабильности.

Включение в геномную ДНК некодирующих последовательностей может приводить и к более специфической защите жизненно важных локусов генома от химических мутагенов. В частности, глобальная защита кодирующих последовательностей от химических мутагенов, поступающих из цитоплазмы в ядро путем радиальной диффузии, могла бы происходить в том случае, если

бы некодирующие последовательности были преимущественно локализованы вблизи поверхности ядер и экранировали последовательности, расположенные ближе к их центральной части. В настоящее время имеются многочисленные экспериментальные данные, указывающие на высокоупорядоченное расположение последовательностей нуклеотидов ДНК в интерфазных ядрах. Значение пространственного расположения отдельных последовательностей ДНК интерфазных хромосом для избирательной, специфической защиты кодирующих последовательностей будет подробнее рассмотрено ниже.

5.3.3. Селективная защита генов от мутаций

Во всех предыдущих рассуждениях речь шла о глобальной защите функционально значимых участков гипотетического генома от спонтанных и индуцируемых мутаций некодирующими последовательностями нуклеотидов. При этом для простоты рассуждений предполагалось, что распределение нуклеотидов геномной ДНК и самих мутагенов в интерфазном ядре гомогенно. В реальном геноме эукариот распределение последовательностей нуклеотидов геномной ДНК в интерфазном ядре далеко не однородно. Достаточно вспомнить, что геном эукариот всегда представлен несколькими хромосомами (в частности диплоидный геном человека заключен в 46 хромосомах), ДНК каждой пары из которых обладает уникальными первичной и пространственной структурами. ДНК индивидуальных хромосом в интерфазном ядре компартиментализована, а плотность упаковки ДНК в различных участках индивидуальных интерфазных хромосом неравномерна.

В этом отношении наиболее изучена и показательна хромомерная организация гигантских политенных хромосом, образующихся в клетках некоторых типов животных и растений. Для таких хромосом в интерфазном ядре характерны различные уровни компактизации хроматина вдоль хроматид, что морфологически проявляется в формировании визуально обнаруживаемых, поперечно расположенных дисков и междисков. В этих хромосомах выделяют три уровня компактизации ДНК. Диска́м (неактивным районам политенных хромосом) свойственен максимальный 100–380-кратный уровень компактизации хроматина по отношению к свободной ДНК. С началом транскрипции, т.е. при переходе участков интерфазных хромосом в активное состояние, уровень компактизации понижается до ~40-кратного, и ДНК

декомпактизуется еще сильнее в условиях максимальной транскрипции, когда молекулы РНК-полимеразы движутся вдоль ДНК одна за другой на небольшом расстоянии друг от друга, как это имеет место, например в кольцах Бальбиани. Уровень упаковки ДНК в хроматине метафазных хромосом самый высокий: исходная длина ДНК в этом случае уменьшается в 6000–7000 раз.

Следует подчеркнуть, что компактизация ДНК вдоль хроматид в интерфазных политенных хромосомах высоко упорядочена в индивидуальных хромосомах генома и обладает абсолютной видовой специфичностью. Это позволяет использовать рисунок поперечной исчерченности индивидуальных политенных хромосом для их идентификации и физического картирования генов. В настоящее время установлено, что некоторые большие гены *Drosophila* (например ген *dunce*, кодирующий фосфодиэстеразу циклического АМР, а также гены *E74* и *Shaker*) локализованы в нескольких дисках и междисках и, следовательно, неравномерно компактизованы по своей длине. Различные уровни упаковки ДНК наблюдают в разных районах одних и тех же политенных хромосом, и для этих уровней, по-видимому, характерны более тонкие градации, чем те, которые выявляются при визуальных наблюдениях в виде поперечной исчерченности политенных хромосом.

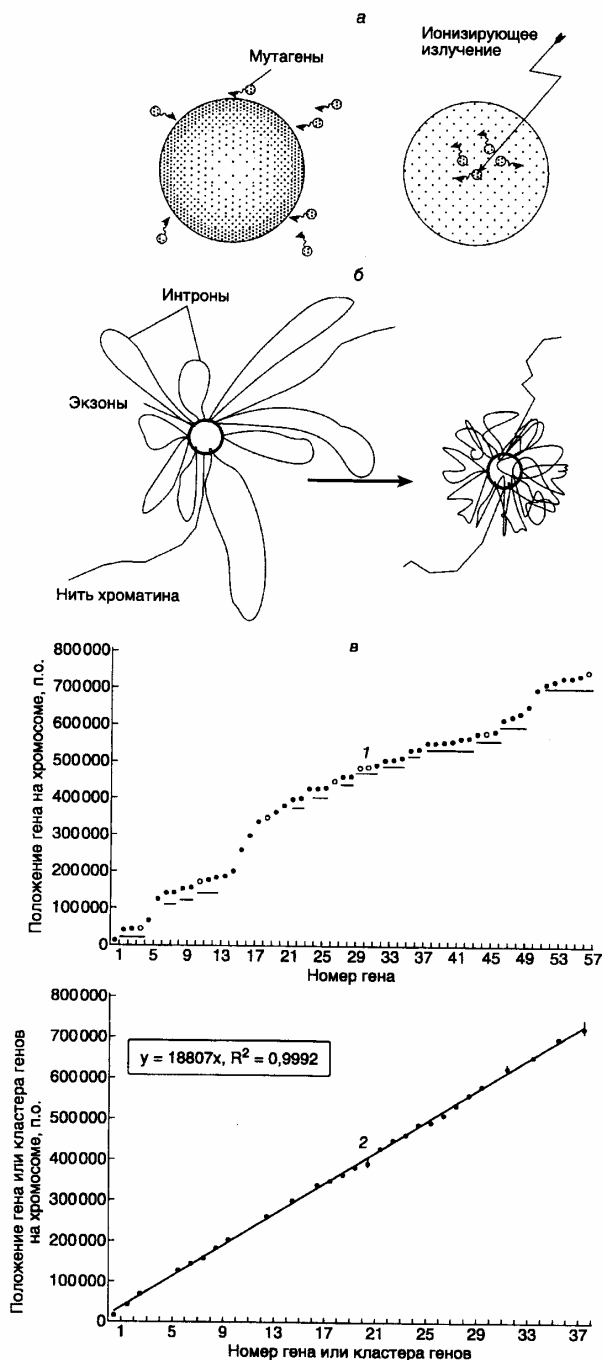


Рис. 1.64. Способы проникновения мутагенов в отдельные микрокомпартменты ядра (а), возможная роль интронов в дифференциальной защите генов (б) и расположение летальных генов на хромосоме 2 дрожжей (в)

а – мутагены, образующиеся вне микрокомпартамента, содержащего ДНК (слева) и внутри него (справа); б – гипотетический ген в виде "розетки" с интронами в виде петель, окружающих объединенные экзоны. Слева – структура гена с расправленными петлями, справа – их гипотетическая реальная пространственная структура, образующая защитную оболочку вокруг экзонов; в – расположение известных жизненно важных генов на хромосоме 2 без предварительной обработки исходных данных (1), а также после обработки с учетом кластеров и гипотетических летальных генов (2). Гены, расстояние между которыми <10 т.п.о., объединяли в кластеры (подчеркнуты), которые на графике 2 изображали в виде отдельных точек с доверительными интервалами, лежащими между первым и последним нуклеотидами кластера. По мере необходимости вводили гипотетические гены (пропуски между точками на графике 2). Указаны коэффициент детерминированности R^2 , описывающий точность аппроксимации полученных данных к линейной функции, и ее уравнение

У большинства хромосом соматических клеток эукариот не происходит полителизации ДНК, и соответствующая часть геномной ДНК заключена в одной хроматиде. Тем не менее, несмотря на то что в интерфазных ядрах большинства соматических клеток визуально не выявляется четкая хромомерная структура индивидуальных хромосом, нет основания полагать, что их хроматиды организованы принципиально иначе, чем в политезированных хромосомах. Действительно, у всех метафазных хромосом соматических клеток животных после их гистохимической окраски выявляется высокоспецифичная поперечная исчерченность (бэндинг), и этот рисунок может быть следствием компактизации высокоупорядоченных хромомерных структур, исходно присутствующих в интерфазных хромосомах. Если такое предположение верно, то у любых интерфазных хромосом соматических клеток эукариот расположение хромомеров вдоль хроматид, а, следовательно, и уровни компактизации ДНК видоспецифичны и являются постоянной характеристикой индивидуальных хромосом и биологических видов в целом.

Рассматривая индивидуальную интерфазную хромосому как потенциальную мишень для мутагенов, можно заключить, что защищенность генетической информации ее индивидуальных локусов от их действия будет прямо пропорциональна уровням компактизации заключенной в них ДНК, т.е. концентрации последовательностей нуклеотидов на единицу ядерного объема, занимаемого этими участками генома. Рассмотрим подробнее процессы, которые могут происходить в микрокомпартаментах (отдельных хромомерах) хромосом при их контакте с химическими внутриядерными мутагенами

(рис. 1.64,а). Как и в случае целых ядер эукариотических клеток, химические мутагены поступают в микрокомпартменты двумя различными путями: образуясь непосредственно внутри них, например под действием ионизирующего излучения, и путем радиальной диффузии через поверхность микрокомпартментов (индивидуальных хромомеров). В первом случае частота мутаций, возникающих на участке заключенной в микрокомпартмент хромосомной ДНК в результате контактов с образующимися здесь же химическими мутагенами, будет обратно пропорциональна концентрации этой ДНК в микрокомпартменте, т.е. уровню ее компактизации в хромомере. Действительно, в продуктивных (приводящих к мутациям) взаимодействиях нуклеотидов ДНК с мутагенами участвует доля нуклеотидов ДНК, пропорциональная количеству нуклеотидов, контактирующих с мутагенами. Эта доля уменьшается при увеличении концентрации нуклеотидов-акцепторов мутагенов в микрокомпартменте. При условии взаимодействия всех k мутагенов с K нуклеотидами рассматриваемого микрокомпартмента (с последующим образованием аддуктов нуклеотид-мутаген и мутаций) частота мутаций F на данном участке ДНК будет равна:

$$F = \frac{k}{K}.$$

В более общем виде это выражение можно записать, как:

$$F = \frac{ak}{K}, \quad (4),$$

где a — коэффициент эффективности взаимодействия мутагенов с нуклеотидами-акцепторами ($0 < a \leq 1$). Если количество нуклеотидов и молекул мутагенов в индивидуальном микрокомпартменте выразить через их концентрации в нем (соответственно C и c), то выражение (4) примет вид:

$$F = \frac{ac}{C},$$

поскольку по определению $k = cV$, а $K = CV$ (V — объем микрокомпартмента, занимаемого индивидуальным хромомером).

Если предположить, что стационарная концентрация мутагенов, образующихся внутри интерфазного ядра эукариотических клеток, одинакова во всех его частях, то частота мутаций, вызываемых такими агентами в индивидуальных хромомерах эукариотических хромосом, будет обратно

пропорциональна концентрации нуклеотидов ДНК этих хромомеров в соответствующих микрокомпартаментах, т.е. обратно пропорциональна плотности упаковки ДНК в хромомере. Следовательно, локальная плотность упаковки цепей ДНК в интерфазном ядре эукариотических клеток может быть важным фактором, определяющим частоту мутаций в соответствующих генетических локусах. При этом необходимо учитывать, что белки являются неотъемлемой частью хроматина эукариотических клеток, и они также могут оказывать влияние на частоту мутаций в индивидуальных хромомерах, затрудняя доступ мутагенов к нуклеотидам ДНК.

Судьба мутагенов, поступающих в ядро эукариотической клетки из цитоплазмы в результате радиальной диффузии, складывается иначе (см. рис. 1.64, а, слева). По-видимому, значительная их часть взаимодействует с нуклеотидами ДНК, расположенной на периферии ядра или внутриядерных хромомеров, и степень защиты как ядра в целом, так и его отдельных микрокомпарментов прямо зависит от плотности такой защитной оболочки. Вторым фактором, регулирующим глубину проникновения мутагенов в ядро и его микрокомпарменты, является реакционноспособность самих мутагенов. Данный параметр, ранее обозначенный как коэффициент эффективности взаимодействия мутагенов с нуклеотидами α (уравнение (4)), определяет время жизни мутагена в ядре, т.е. как скоро молекула мутагена прореагирует с потенциальными нуклеотидами-акцепторами, которые мутаген встречает в процессе диффузии внутри ядра. Третьим фактором, от которого зависит эффективность защиты кодирующих жизненно важных последовательностей нуклеотидов генома от мутаций, вызываемых такими мутагенами, является пространственное расположение самих внутриядерных последовательностей. Очевидно, что вероятность встречи с мутагенами и взаимодействия с ними больше у нуклеотидов последовательностей, локализованных на периферии ядра или хроматиновых доменов – хромомеров. С учетом этого фактора можно полагать, что максимальная защита имела бы место в том случае, если бы мутагены первыми встречали избыточные некодирующие последовательности, мутации в которых нейтральны в отношении влияния на жизнеспособность клеток многоклеточного эукариотического организма.

В последнее время появляется все больше экспериментальных данных, свидетельствующих о высокоупорядоченном распределении специфических

последовательностей нуклеотидов в интерфазных ядрах эукариот. Одна группа данных такого рода, связанных с упорядоченной конденсацией нитей ДНК вдоль хроматид индивидуальных интерфазных хромосом, была уже обсуждена выше. Помимо этого, активно рассматривается модель петельно-доменной организации генов в эукариотических хромосомах. Неслучайное распределение MAR- и SAR-последовательностей в геномной ДНК эукариот и их ассоциация с белками ядерного матрикса (скэффолда) в интерфазе клеточного цикла, по-видимому, весьма специфически контролируют пространственное расположение протяженных участков интерфазных хромосом в ядрах. Кроме того, на более низком (хромомерном) уровне компактизации интерфазных хромосом также обнаружены специфические способы укладки ДНК в виде больших и малых петель, вероятно, относящихся к хромомерам и "розеткам" соответственно (см. рис. 1.64,б). Не исключена возможность, что в соответствии с моделью Н.А. Резника и соавторов (1991 г.) образование розеток из ДНП-петель специфически отражает интрон-экзонную структуру конкретных генов. Если такая модель соответствует действительности, то петли интронов генов, заключенных в розетки, могут создавать специфическую защитную оболочку, предохраняющую экзоны генов от контакта с мутагенами. В этой связи соотношение суммарных размеров интронов и экзонов индивидуальных генов может оказаться существенным параметром, характеризующим уровень защищенности функционально значимых участков генов от химического мутагенеза. Чем больше отношение суммарной длины интронов к суммарной длине экзонов в конкретном гене, тем более плотную защитную оболочку интроны могли бы сформировать вокруг компактизованных экзонов. Такое соотношение может указывать на уровень защищенности конкретных генов (или даже их частей) от химических мутагенов.

Описанные выше структурные особенности организации генетического материала позволяют предполагать разную и генетически детерминированную доступность индивидуальных участков интерфазных хромосом для химических мутагенов. Имеются, по крайней мере, три группы экспериментальных данных, указывающих на то, что различные участки одного и того же генома спонтанно изменяются с помощью мутаций с разной скоростью. Прежде всего, это прямые экспериментальные указания на неодинаковую мутабельность как различных генов одного и того же генотипа, так и одного и того же гена в генотипах разных

биологических видов.

При другом подходе к исследованию этой проблемы путем компьютерного анализа баз данных последовательностей нуклеотидов было установлено, что степень эволюционной дивергированности конкретных участков генома у разных видов млекопитающих непостоянна и может изменяться от локуса к локусу. Скорость накопления замен нуклеотидов в конкретных локусах приблизительно одинакова как для кодирующих последовательностей нуклеотидов генов (экзонов), так и для соседствующих с ними некодирующих последовательностей, казалось бы, не подверженных давлению отбора. Этот, на первый взгляд, странный результат вполне естественно объясняется в рамках предлагаемой модели дифференциальной защиты различных хромосомных локусов от спонтанных мутаций как следствие упорядоченного гетерогенного распределения геномной ДНК внутри интерфазных ядер. Действительно, рядом расположенные последовательности должны быть одинаково защищены от мутагенов независимо от их функциональной нагрузки. Такой механизм защиты конкретных генетических локусов от спонтанных и индуцированных мутаций мог бы приводить к мозаичной эволюции генома, предполагаемой в модели Б.Ф.Купа.

Наличие в геноме филогенетически консервативных последовательностей нуклеотидов, как правило, принято объяснять жестким давлением отбора на эти последовательности в процессе эволюции. Например, к числу наиболее консервативных из известных белков относятся гистоны H3 и H4, у которых одна замена на 100 аминокислотных остатков, по-видимому, эволюционно фиксируется в полипептидной цепи не чаще чем в несколько сотен миллионов лет. На противоположном полюсе находятся иммуноглобулины и фибринопептиды, у которых такие мутационные изменения фиксируются в геноме в сотни раз быстрее. В соответствии с традиционной точкой зрения медленно и быстро эволюционирующие белки различаются по содержанию в них функционально значимых сайтов, изменение которых под действием мутаций несовместимо с жизнью. Можно предполагать, однако, что селективная консервативность их генов обеспечивается и структурными особенностями соответствующих локусов генома конкретных видов организмов. Функционирование такого механизма предотвращало бы тяжелые последствия в виде летальных исходов, которыми приходится расплачиваться организмам и

соматическим клеткам за мутации в соответствующих локусах, если бы те возникали с такой же частотой, что и в функционально незначимых частях генома. Подобный механизм мог бы быть эффективным в селективной защите от мутаций генетических локусов, содержащих множественные копии генов, так как избыточность генетической информации создает особенно благоприятную почву для дивергенции гомологичных последовательностей. Действительно, дубликации генов и их последующая дивергенция рассматриваются в настоящее время как ключевые механизмы эволюционного процесса.

Еще одна группа данных, свидетельствующих в пользу модели дифференциальной защиты отдельных генетических локусов от спонтанных мутаций, была получена при анализе скоростей образования мутаций в трансгенах, интегрированных в различные участки генома организма-хозяина. Значительную вариабельность в скоростях мутирования локусов мини- и микросателлитов, обнаруженную в геноме млекопитающих и проявляющуюся в различиях уровней их внутривидового полиморфизма, также частично можно объяснить с помощью обсуждаемой модели.

Таким образом, видоспецифическое пространственное расположение ДНК отдельных генетических локусов интерфазных хромосом создает материальную основу для их дифференциальной чувствительности по отношению к химическим мутагенам экзогенного и эндогенного происхождения. Избирательная доступность индивидуальных нуклеотидов ДНК таких локусов может детерминировать относительную частоту спонтанных мутаций в отдельных локусах геномной ДНК эукариот. Все это может иметь большое значение для онтогенетического и филогенетического развития биологических видов.

5.3.4. Высокоупорядоченное расположение летальных генов на хромосомах

Если гипотеза о наличии внутри ядер генетически детерминированных, пространственно упорядоченных участков геномной ДНК является верной, то это влечет за собой важное следствие. В этом случае в процессе эволюционных преобразований геномов их участки, максимально защищенные от действия химических мутагенов, в результате транслокаций и других хромосомных перестроек могли быть заняты генами, функциональная целостность которых особенно важна для выживания клеток или организмов.

Таковыми генами являются, прежде всего, жизненно важные (летальные) гены, инаktivация которых под действием мутаций приводит к гибели клеток. Действительно, организмы, у которых критические гены максимально защищены от действия мутагенов, должны были бы обладать эволюционным преимуществом перед организмами, геном которых не обладает этими свойствами, и в первую очередь сохраняться естественным отбором. Если это так, то летальные гены маркируют в геноме современных организмов участки, максимально благоприятные для их целостности, т.е. наиболее защищенные от действия мутагенов, наиболее доступные для белков системы репарации и т.п. В этой связи можно предполагать, что линейное расположение жизненно важных генов вдоль ДНК отражает особенности пространственной организации ДНК в ядрах эукариот на высшем уровне.

Используя эти рассуждения в качестве рабочей гипотезы, мы исследовали расположение жизненно важных генов на всех 16 хромосомах дрожжей *Saccharomyces cerevisiae*, первичная структура которых в настоящее время полностью определена. Для этого объекта точно известно положение на хромосомах многих жизненно важных генов. В ходе исследований было установлено, что на всех хромосомах дрожжей имеются участки, в которых известные на момент исследования летальные гены (или их кластеры) расположены периодически (на равном расстоянии друг от друга). В качестве примера рассмотрим положение жизненно важных генов на одной из хромосом.

Хромосома 2 является одной из пяти наиболее крупных хромосом дрожжей, длина которой составляет ~813 т.п.о. Хромосома содержит 429 открытых рамок считывания (ОРС), среди которых ко времени проведения анализа 57 были определены как жизненно важные гены, 193 – как нелетальные, а указанные свойства остальных 179 ОРС в настоящее время неизвестны. Следовательно, количество известных нежизненно важных генов хромосомы 2 приблизительно в 3 раза превышает число известных жизненно важных, и это соотношение к настоящему времени соблюдается у всех хромосом дрожжей.

Летальные гены расположены на хромосоме неравномерно. Имеются места их скопления, особенно ярко выраженные в центральной части хромосомы, а также участки, в которых они встречаются не так часто (см. рис. I.64, в, 1). Даже без предварительной обработки данных можно было видеть

дискретность расположения летальных генов и их кластеров, в которые объединяли гены, если расстояние между ними было меньше 10 т.п.о. (подчеркнуты на рисунке). После представления каждой группы генов отдельного кластера в виде одной точки (с доверительными интервалами, расположенными между началами и концами кластеров) и введения в график нескольких свободных мест для гипотетических, пока неизвестных генов (или их кластеров) периодический характер распределения кластеров летальных генов на хромосоме 2 становится очевидным (см. рис. 1.64, в, 2). Видно, что места предпочтительного расположения летальных генов или их кластеров расположены на хромосоме 2 периодически. При этом период расположения генов или их кластеров (повторяющееся равное расстояние между следующими друг за другом генами) численно равен выраженному в парах оснований угловому коэффициенту линейной функции, которая описывает последовательность координат генов на графиках.

Таким образом, после минимальных преобразований исходных данных, потребовавших введения в общей сложности девяти гипотетических генов или их кластеров (места, в которых отсутствуют точки на графике), а также выделения 14 групп кластеров генов, последовательность координат летальных генов и их кластеров удовлетворительно описывается линейной функцией с угловым коэффициентом 18 807 ($R^2 = 0,9992$). Это указывает на периодический характер расположения летальных генов вдоль всей хромосомы 2 с периодом $\sim 18,8$ т.п.о.

Разработанная нами методика анализа летальных генов, кратко рассмотренная на примере хромосомы 2, была использована для исследования остальных хромосом *S. cerevisiae*. Оказалось, что на каждой из исследованных хромосом можно выделить участки с периодическим расположением летальных генов. По характеру распределения жизненно важных генов исследованные хромосомы разделяются на две группы. В одной из них (хромосомы 1, 2, 5–7, 9, 11, 12 и 16) летальные гены (и их кластеры) образуют непрерывную последовательность и распределены равномерно по всей длине хромосом. При этом у семи из девяти хромосом этой группы периоды расположения генов очень близки и лежат в пределах 22,0–25,8 т.п.о.

Ко второй группе относятся хромосомы, на которых летальные гены образуют несколько участков с разной периодичностью. Для относящихся к

этой группе хромосом 2, 10 и 14 характерны два участка периодически расположенных летальных генов. При этом участки хромосомы 3, обладающие очень близкими периодами, локализованы симметрично относительно центра хромосомы, в окрестностях которого летальные гены пока не обнаружены. Равные по своим размерам хромосомы 10 и 14 обладают похожим строением в отношении анализируемого признака: за участком с меньшей периодичностью следует область хромосомы, на которой расстояние между летальными генами или их кластерами больше.

На хромосомах 4, 8 и 13 можно выделить по три области с периодически расположенными летальными генами, структура которых очень похожа. В этом случае области с меньшими периодами фланкируют участки хромосом, на которых расстояния между периодически расположенными летальными генами больше. Интересно, что на хромосоме 8 близкие по периодичности концевые участки (периоды 22,3 и 23,7 т.п.о.) расположены симметрично относительно центральной области хромосомы, для которой характерен приблизительно вдвое больший период чередования жизненно важных генов (40,9 т.п.о.). Не исключено, что по мере открытия новых летальных генов в центральной части этой хромосомы все три участка сольются в единую область, периодичность расположения летальных генов в которой будет близка таковой в отмеченных концевых участках (22–23 т.п.о.).

Для крупной хромосомы 15 характерно наличие пяти участков с периодически расположенными летальными генами. И на этот раз периоды участков 1, 3 и 5 очень близки (20,5; 23,0 и 23,9 т.п.о. соответственно). В то же время расстояния между периодически расположенными жизненно важными генами этой хромосомы на участке 2 приблизительно вдвое больше (46,2 т.п.о.).

Несмотря на то что точные числовые значения периодов в расположении летальных генов различаются как между хромосомами, так и между конкретными участками индивидуальных хромосом дрожжей, складывается впечатление об универсальном характере расположения анализируемых летальных генов. Действительно, среди 29 обнаруженных участков с периодическим расположением летальных генов на 16 исследованных хромосомах у 17 из них значения периодов лежат в пределах 19,7–25,8 т.п.о., а из оставшихся 12, по крайней мере, три значения могут рассматриваться как

кратные им, т.е. они подтверждают ту же структурную закономерность. Обнаруженная периодичность в расположении жизненно важных генов хромосом дрожжей может указывать на наличие в интерфазных ядрах дрожжей периодически повторяющихся пространственных структур хроматина высокого порядка, что может создавать особые биохимические условия для находящихся в них генов, например иметь отношение к защите жизненно важных генов от мутаций, как физически, так и обеспечивая эффективное функционирование ферментов системы репарации.

Периодичность расположения жизненно важных генов на хромосомах дрожжей как возможное отражение пространственной структуры хроматина. Существование информационных макромолекул, особенно таких гигантских, как молекулы ДНК, полностью зависит от их упорядоченной пространственной структуры. Последовательные циклы компактизации и декомпактизации хроматина сопровождают каждое деление эукариотических клеток и являются одним из самых универсальных и распространенных генетических процессов в живой природе. Точность и эффективность этого процесса очень высоки. Если исходная пространственная структура хроматина в интерфазных ядрах еще во многом остается непонятной, то структура метафазных хромосом, выявляемая на цитогенетическом уровне, является консервативным видоспецифическим признаком. Последние модели строения метафазных хромосом указывают на наличие у них центрального остова, включающего в себя тандемно повторяющиеся MAR/SAR-последовательности, и упорядоченных боковых петель хроматина. Трудно представить себе, чтобы имеющиеся связи между участками хроматина, сближенными в метафазных хромосомах, полностью утрачивались при его декомпактизации в интерфазе клеточного цикла, поскольку это должно было бы затруднять и замедлять его циклически повторяющуюся сборку в начале каждого митоза. Одним из указаний на сохранение таких связей является наличие в интерфазных ядрах особых хромосомных зон, занимаемых индивидуальными декомпактизованными хромосомами, которые не перемешиваются друг с другом.

Основа пространственной упорядоченности ДНК в составе хроматина заложена в ее первичной структуре. Как известно, монотонно следующие друг за другом четыре азотистых основания ДНК образуют правильную двойную

спираль, шаг которой в случае В-формы ДНК составляет 10,5 нуклеотидов на виток двойной спирали. Именно такая монотонная организация молекулы ДНК в конечном итоге дает возможность формироваться на ней, как на матрице и прямом участнике процесса, периодически повторяющихся нуклеосом. На этом первом уровне пространственной организации хроматина так называемые коровые частицы нуклеосом (тетрамер гистонов H3/H4, фланкированный димерами гистонов H2A/H2B, с закрученным вокруг них участком ДНК длиной в 146 п.о.) разделены участками линкерной ДНК длиной ~50 п.о. Особенности пространственной структуры хроматина на более высоких уровнях (соленоид и петельно-доменный уровни компактизации хроматина) до конца не ясны и по-прежнему остаются предметом дискуссий. Имеющиеся экспериментальные данные указывают на существование периодически повторяющихся пространственных структур и на высших уровнях упаковки интерфазного хроматина.

Ограниченная инкубация нативного хроматина животных и растений с нуклеазами позволяет обнаруживать с помощью электрофореза в импульсном электрическом поле образование дискретных фрагментов ДНК двух классов: крупных, длиной ~300 т.п.о. и более коротких – ~50 т.п.о. Использование топоизомеразы для расщепления ДНК в основаниях петель, ассоциированных с ядерным матриксом, приводит к накоплению фрагментов ДНК приблизительно того же размера. Наконец, деградация ДНК на ранних стадиях апоптоза начинается с образования аналогичных по размерам фрагментов геномной ДНК. К сожалению, соответствующие данные относительно пространственной организации хроматина дрожжей мне неизвестны. Помимо выше отмеченных факторов материальной основной формирования периодически организованных пространственных структур хроматина могут быть и повторяющиеся последовательности нуклеотидов, в большом количестве встречающиеся в геноме высших эукариот.

Обнаруженное в ходе нашего исследования периодическое распределение летальных генов вдоль всех 16 хромосом дрожжей по-своему указывает на наличие возможной связи между пространственной структурой их хроматина и функциональной значимостью генетического материала, включенного в соответствующие последовательности. Как уже отмечалось выше, значения большинства периодов между летальными генами и/или их

кластерами лежат в пределах 20–26 т.п.о., что представляет собой величину того же порядка, что и размеры фрагментов ДНК, образующихся при ограниченном нуклеазном гидролизе нативного хроматина. На мой взгляд, жизненно важные гены дрожжей маркируют места хромосом, наиболее безопасные для их существования. Поскольку наибольшую опасность для клетки представляют мутации, инактивирующие их летальные гены, места их расположения могут быть в большей степени защищены от действия химических мутагенов, с которыми организм в избытке сталкивается в процессе своей жизнедеятельности. Такую защиту для последовательностей нуклеотидов могли бы обеспечивать, например внутренние части хроматиновых глобул. Действительно, уже сам факт наличия у нативного хроматина упорядоченно расположенных участков ДНК, в большей или меньшей степени защищенных от действия нуклеаз (что и дает возможность образования характерных дискретных фрагментов ДНК), однозначно указывает на существование в хроматине последовательностей нуклеотидов, по-разному защищенных от действия мутагенов. В соответствии с вышеизложенным мы предполагаем наличие вдоль хромосом дрожжей периодически повторяющихся мест с большей или меньшей защищенностью от действия мутагенов. В зависимости от тонкой пространственной структуры хроматина в этих участках уровни защищенности последовательностей нуклеотидов от мутагенов могут варьировать от участка к участку и приводить к генетической детерминированности скоростей спонтанного мутагенеза в конкретных генетических локусах.

Возможное биологическое значение обнаруженной периодичности расположения летальных генов на хромосомах дрожжей. Хроматин всех эукариот построен в общих чертах одинаково, поэтому обнаруженные у дрожжей особенности его строения и вытекающие из этого следствия хотелось бы рассмотреть применительно и к геному высших эукариот. В этой связи необходимо обратить внимание на четыре момента развиваемой концепции, которые могут иметь общебиологическое значение.

1. Значительно большее разнообразие последовательностей нуклеотидов различных типов, присутствующих в геноме высших эукариот, по сравнению с геномом дрожжей, создает условия для формирования более разнообразных и богатых в функциональном отношении пространственных

структур хроматина в интерфазных ядрах. Наличие же таких структур, в свою очередь, предполагает существование у высших эукариот и более тонкого контроля скоростей спонтанного мутагенеза в конкретных генетических локусах по обсуждаемому механизму. Такая разная генетическая детерминированность темпов мутационных изменений генетических локусов у биологических видов могла бы контролировать направление их возможных эволюционных преобразований и историческое развитие таксонов.

2. Наличие в интерфазных ядрах эукариот участков ДНК, максимально защищенных от действия мутагенов, предполагает существование в них и генетических локусов с минимальным уровнем защиты. Из этого следует, что делеции или вставки в окрестностях защищенных локусов могут сдвигать генетические локусы в менее благоприятное, с точки зрения защиты, положение на хромосоме. Для менее защищенных локусов могут возникать обратные ситуации. Следовательно, делеции или вставки, а также природный геномный полиморфизм (в частности обнаруживаемый по длинам рестрикционных фрагментов ДНК) могут быть причиной (и новым механизмом) возникновения локального мутаторного фенотипа в соматических клетках высших организмов. Известно, что мутаторный фенотип часто предшествует малигнизации клеток и сопровождает рост опухолей. Если предполагаемый нами механизм функционирует, то возникновение делеции или вставки по соседству с критическим локусом (протоонкогеном или антионкогеном), контролирующим развитие заболевания, должно переводить мутантный организм в группу риска с повышенной вероятностью изменения этого локуса под действием мутаций. Другим примером могла бы быть лейденская мутация в факторе V системы свертывания крови, ассоциированная с развитием тромбозов, которая чрезвычайно широко распространена в европейской популяции и не встречается у ориенталов. Это и другие подобные явления можно объяснить неблагоприятным пространственным расположением соответствующего генетического локуса в интерфазном ядре у индивидуумов европейской популяции, принадлежащих к группе риска, которое делает локус легко доступным для химических мутагенов или препятствует функционированию ферментов системы репарации.

3. Исходя из всего вышеизложенного, можно полагать, что "избыточные" последовательности генома эукариот обеспечивают

необходимую пространственную структуру хроматина в интерфазных ядрах, создавая оптимальные условия для генов с точки зрения их экспрессии и защиты от мутационных изменений, что является жизненно важным фактором существования эукариот.

4. Интроны в генах эукариот могут обеспечивать специфическое пространственное расположение экзонов в интерфазных ядрах, оптимальное с точки зрения их защиты от мутаций и эффективности экспрессии генов. Например, делеционное удаление интронов из гена может приводить к тому, что его 5'-концевая регуляторная часть в процессе упаковки ДНК в хроматин попадет внутрь хроматиновой глобулы и станет недоступной РНК-полимеразам, факторам транскрипции и другим регуляторным белкам. В такой ситуации наиболее важными для гена становятся длины его интронов, а не их первичная структура.

5.3.5. Возможный смысл парадокса С

У организмов, находящихся на примерно одинаковых ступенях эволюционного развития, часто наблюдаются значительные вариации в размерах геномов (см. главу 1). Например, у некоторых видов рыб, относящихся к разным отрядам и подклассам, размеры геномов заметно различаются. Наименьшие геномы характерны для некоторых неродственных видов костистых рыб. В частности, у малоротой корюшки, меченосца или камбалы размер генома приблизительно в 5 раз меньше генома млекопитающих. В то же время у двоякодышащих рыб (одни из немногих выживших представителей кистеперых рыб, которые морфологически практически не изменились за миллионы лет своего существования) размер генома, по крайней мере, в 35 раз превышает размер генома плацентарных млекопитающих. Ввиду большого морфологического и физиологического сходства этих видов рыб можно предположить, что различия в размерах их геномов, главным образом, определяются относительным содержанием в них избыточных последовательностей нуклеотидов.

Развивая концепцию о стабилизирующем влиянии избыточных последовательностей нуклеотидов ДНК на генетическую информацию, заключенную в геноме многоклеточных организмов, можно предполагать, что различия в размерах геномов у близких видов многоклеточных организмов

отражают существенные особенности в функционировании их генетического аппарата и его внутриклеточного окружения. У организмов с высоким содержанием избыточной ДНК могут менее эффективно функционировать ферменты репаративной системы, что, в свою очередь, усиливает (имитирует) экзогенное и эндогенное мутагенные воздействия. Действительно, эффективность функционирования систем эксцизионной репарации, по-видимому, существенно различается даже в клетках разных видов млекопитающих. Кроме того, у таких организмов могло бы быть более высоким внутриядерное содержание эндогенных мутагенов в силу видовых особенностей их метаболизма. В связи с этим весьма вероятно, что гигантский размер геномов двоякодышащих рыб и амфибий отражает одну или несколько таких особенностей функционирования их генетического аппарата и его внутриклеточного окружения. Эволюционное включение в их геном большого количества защитных избыточных последовательностей нуклеотидов могло значительно стабилизировать геном и, по-видимому, позволило этим организмам пройти морфологически неизменными через миллионы лет своего существования. С другой стороны, малый размер генома других видов может сочетаться с более эффективно работающими системами репликации и репарации, а также способствовать ускоренному видообразованию.

Таким образом, повышение точности функционирования систем репликации и репарации ДНК, с одной стороны, и увеличение размера генома за счет включения в него некодирующих последовательностей нуклеотидов, с другой, могут приводить к одному и тому же эволюционному последствию: увеличению информационной стабильности генома. Поэтому размер генома современных эукариот эволюционно оптимизирован в отношении максимально допустимой частоты мутаций, совместимых с жизнеспособностью конкретных биологических видов. Сохранение видоспецифических соотношений между кодирующими и некодирующими последовательностями генома эукариот может быть следствием естественного отбора, отсекающего крайние варианты, у которых или слишком мало, или чрезмерный избыток некодирующих последовательностей.

Эволюционное увеличение размера генома понижает требования таких видов к точности функционирования систем репарации геномной ДНК. В результате снижение давления отбора на эти ферментативные системы могло

способствовать накоплению в них мутаций, уменьшающих точность функционирования таких систем. Следовательно, суммарный размер генома эукариотического организма отражает не только потребность организма в определенном количестве генетической информации для обеспечения соответствующего уровня сложности его биологической организации, но и особенности жизнедеятельности организма, связанные с интенсивностью экзогенных и эндогенных мутагенных воздействий.

Подводя итоги вышесказанному, необходимо еще раз отметить, что, по крайней мере, две особенности строения генома эукариот могут оказывать влияние на частоту мутаций, возникающих в нем в процессе репликации под действием экзогенных и эндогенных мутагенов. Во-первых, включение избыточных последовательностей нуклеотидов приводит к глобальной защите всех функционально значимых последовательностей генома от эндогенных и экзогенных мутагенов. Избыточные последовательности нуклеотидов генома эукариот весьма существенно стабилизируют геном, что, возможно, является необходимым и достаточным условием для эволюционного появления многоклеточности. Во-вторых, внутриядерная компартментализация последовательностей нуклеотидов геномной ДНК, при которой происходит специфическая упаковка нитей хроматина в индивидуальных компартментах, занимаемых хромомерами и более крупными блоками последовательностей нуклеотидов, должна также сопровождаться изменением частоты мутаций в конкретных генетических локусах пропорционально локальной внутриядерной концентрации ДНК и пространственному расположению отдельных генетических локусов друг относительно друга.

Эволюционно сложившиеся отношения между суммарными длинами экзонов и интронов в индивидуальных генах, а также уровни упаковки и пространственное расположение ДНК в отдельных генетических локусах могли бы указывать на тот максимально допустимый темп мутационных изменений экзонов, который совместим с жизнеспособностью организмов в онтогенезе. С другой стороны, эти соотношения были бы своеобразной генетической программой, предопределяющей и филогенетическое развитие видов. Действительно, селективная видоспецифическая защита отдельных генетических локусов от спонтанного и индуцированного мутагенеза должна сопровождаться преимущественным образованием мутаций в локусах,

наименее защищенных некодирующими последовательностями нуклеотидов, на фоне которых и разворачиваются основные события, связанные с естественным отбором. При этом уровни защищенности отдельных генетических локусов определяют различные темпы изменений этих локусов в филогенезе разных таксономических групп организмов. Частоты спонтанных мутаций в индивидуальных генах ограничиваются достаточно узкими рамками, определяемыми пространственной структурой как самих генов, так и более протяженных генетических локусов, включающих в себя некодирующие последовательности нуклеотидов. Такие ограничения могут быть преодолены при более мощных мутагенных воздействиях, однако соотношение частот мутаций, возникающих в разных частях генов, в основном, должно сохраняться, в том числе из-за разной доступности этих частей химическим мутагенам.

Все вышеперечисленные причины могут в конечном счете определять дискретный популяционный полиморфизм и направление изменчивости фенотипов индивидуальных биологических видов, наблюдаемые в природе. В этих терминах можно было бы объяснить закон гомологических рядов Н.И. Вавилова, в соответствии с которым у родственных видов, родов и даже семейств организмов наблюдаются сходные ряды фенотипической изменчивости. Действительно, данное явление вполне естественно объясняется общностью генотипов таких таксономических групп и общностью пространственной структуры их геномов, определяющей мутабельность отдельных генетических локусов и направление эволюционных преобразований этих генов и признаков. Такая генетически детерминированная изменчивость генотипов индивидуальных биологических видов и может дать гомологические ряды фенотипических признаков родственных организмов, которые формируются направленно изменяющимися генотипами. В то же время для образования признаков, выводящих организмы из их таксономических групп, требуются более радикальные преобразования генотипов, чем точковые мутации, которые бы открывали новые участки генома для интенсивного спонтанного мутагенеза. Направленная изменчивость генотипов, определяемая пространственной структурой и составом интерфазных хромосом, должна допускать в больших популяциях одновременное образование одних и тех же мутантных фенотипических признаков у большого числа особей и ускорять процесс видообразования.

Не исключено, что еще большее, чем интроны, отношение к генетической программе филогенетического развития имеют крупные блоки повторяющихся последовательностей индивидуальных хромосом, которые окружают и в разной степени защищают от мутационных изменений участки генома, наиболее важные для видообразования и сохранения вида как такового. Анализ и картирование этих блоков в геноме многоклеточных организмов могут способствовать выявлению новых функционально значимых участков генома и их экспериментальному исследованию. Не менее интересные результаты может принести и анализ интрон-экзонной структуры известных генов, а также фланкирующих генов избыточных последовательностей по уровню защищенности индивидуальных генов от мутаций. Такой анализ может по-новому осветить генетическую значимость уже известных участков генома. С использованием аналогичного подхода могут быть выявлены новые жизненно важные гены, которые организм особенно бережно укрывает от мутаций избыточными последовательностями нуклеотидов.

Организовав геном эукариот таким изящным (но не безупречным) образом, природа сама указывает на его слабые места и возможные пути дальнейшего совершенствования. Уязвимость современного генома в отношении неблагоприятных экологических факторов (в первую очередь, антропогенных) может быть связана с тем, что в доисторические времена, когда происходила адаптивная эволюция организмов, экологическая обстановка была существенно благоприятнее. Крупные геномы гораздо более чувствительны как мишени для ионизирующих излучений, а этот фактор в те времена мог не иметь большого значения, и эволюционирующим организмам не нужно было к нему приспосабливаться. Уязвимость такого сложноорганизованного генома, как геном эукариот, заключается еще и в том, что его мутационные изменения, нарушающие пространственную организацию укладки ДНК в интерфазных ядрах, могут быть причиной возникновения мутаторного фенотипа мутантных соматических клеток. Этот фенотип, в свою очередь, может вызывать тяжелые патологические изменения организма, включая онкологические, аутоиммунные и другие тяжелые заболевания. С другой стороны, введение генно-инженерными методами генетически нейтральных последовательностей нуклеотидов, изменяющих доступность для мутагенов особенно важных генетических локусов генома, могло бы повысить

устойчивость организма к мутагенным воздействиям, тем самым оградив его от многих патологий и увеличив продолжительность жизни. Такой генно-инженерный подход мог бы стать одним из элементов превентивной генотерапии.

С учетом всего сказанного вряд ли можно считать избыточные последовательности нуклеотидов эукариот "эгоистичными" геномными паразитами, как это часто обсуждается в современной литературе. Избыточные последовательности нуклеотидов в геноме, количество и локализация которых, по-видимому, строго сбалансированы по отношению к функционально значимым последовательностям в процессе эволюции, ведут себя вполне "альтруистично", принимая удар мутагенов на себя и специфически защищая жизненно важные участки ДНК от мутаций. Их "самоотверженное" "альтруистичное" поведение способствует сохранению хрупкой стабильности генома эукариот на приемлемом уровне и дает возможность многоклеточным организмам существовать в виде гигантских клонов высококооперированных соматических клеток, а также, вероятно, перспективу и вектор их филогенетического развития.

ГЛАВА 6. СОВРЕМЕННАЯ КОНЦЕПЦИЯ ГЕНА

Подводя итоги рассмотрению основных достижений в исследованиях структуры и механизмов функционирования генов, представляется целесообразным более подробно обсудить вопрос о том, что же такое ген в свете современных экспериментальных данных.

Классическое представление о гене как цистроне, а также модель Ф. Жакоба и Ж. Моно регуляции активности полицистронных оперонов бактерий явились вершиной развития молекулярной генетики 1960-х годов. Работы С. Бензера по изучению тонкой структуры гена *gII*-области бактериофага Т4 продемонстрировали предел (а точнее – беспредельность) разрешающей способности генетического анализа в применении к бактериофагам и одновременно невозможность получения таких же точных результатов при его использовании для исследования сложноустроенных организмов с более продолжительным жизненным циклом. К началу 1970-х гг. в отношении структуры и функционирования гена, по крайней мере прокариотического, складывалось вполне определенное и логичное мнение о том, что ген представляет собой непрерывную последовательность нуклеотидов ДНК, состоящую из двух частей – регуляторной и структурной. В модели гена 1960-х годов регуляторная часть всегда предшествует структурной и определяет уровень его экспрессии через взаимодействие с регуляторными белками. Структурная часть гена кодирует единственную полипептидную цепь белка или молекулу РНК и колинеарна белковому продукту, т.е. последовательность нуклеотидов гена однозначно определяет и соответствует последовательности аминокислот в кодируемом этим геном белке. Модель гена 1960-х годов предусматривает его четкие границы и постоянную локализацию на хромосоме. Изменения в первичной структуре гена и его перемещения в геноме могли быть только следствием непредсказуемых мутаций и должны были приводить к нарушению его функционирования. Таковы общие представления о гене, которые складывались в молекулярной генетике на основании экспериментальных данных до разработки методов генной инженерии. Последние 20 лет интенсивного развития молекулярной биологии и генетики с широким использованием методов рекомбинантных ДНК изменили эти

представления. Сегодня совершенно ясно, что ген не всегда колинеарен РНК или белку, которые закодированы в последовательности его нуклеотидов – одна и та же последовательность нуклеотидов ДНК может кодировать разные полипептидные цепи, а нестабильность генов может быть генетически запрограммированной. Прежде чем сделать попытку сформулировать современное определение термина "ген", обобщим основные свойства генов, подробно рассмотренные в первых пяти главах книги.

Ген не всегда колинеарен своим РНК или белкам. Большинство генов эукариот обладает мозаичной структурой кодирующих и не кодирующих последовательностей нуклеотидов – экзонов и интронов. Лишь в результате сплайсинга предшественников мРНК заключенные в них интроны вырезаются и восстанавливается колинеарность гена, мРНК и белкового продукта (см. раздел 2.2.3). Удаление избыточных последовательностей в результате сплайсинга происходит не только на уровне РНК, но и на уровне полипептидных цепей. Недавно обнаруженный сплайсинг белков у дрожжей *S. cerevisiae* и других микроорганизмов (см. раздел 3.6.4) по-новому осветил возможности механизмов, контролирующих реализацию генетической информации.

Не менее удивительными, с привычной точки зрения, являются и механизмы редактирования кодирующего потенциала РНК на посттранскрипционном уровне. Расширение исследований в этой области показывает, что такое редактирование РНК широко распространено у животных, растений и микроорганизмов. Факты корректировки генетической информации на посттранскрипционном уровне оказались неожиданными. Почему организму выгоднее не сохранять всю необходимую генетическую информацию в самом исходном гене, а корректировать ее уже в процессе реализации? К сожалению, пока на этот вопрос нет однозначного ответа. Однако можно полагать, что такая корректировка кодирующего потенциала РНК имеет непосредственное отношение к дополнительной стабилизации генетической информации в наиболее уязвимых для мутагенеза генетических локусах (см. раздел 2.2.2). Если мое объяснение верно, то существование механизма редактирования первичной структуры РНК на посттранскрипционном уровне не изменяет классических представлений о функционировании генов.

Одна и та же последовательность нуклеотидов ДНК может кодировать разные полипептидные цепи. Генетический материал многих вирусов и бактериофагов организован компактно. Например, на определенных участках генома колифага f2 или бактериофага ϕ X174 для кодирования белка используются две возможные рамки считывания. Уникальное сверхкомпактное кодирование генетической информации характерно для мобильного генетического элемента бактерий IS5, у которого одна из цепей ДНК содержит два перекрывающихся гена, а комплементарный им участок другой цепи включает в себе третий ген. В таких случаях можно, по-видимому, говорить о наличии в одной и той же последовательности нуклеотидов нескольких генов в классическом смысле этого слова.

Кодирование разных полипептидов одной и той же последовательностью нуклеотидов ДНК имеет место и у эукариот, а также их вирусов. Так, из одного предшественника мРНК в результате альтернативного сплайсинга образуется несколько зрелых мРНК, кодирующих разные белки (см. раздел 3.3.2). Следовательно, один и тот же ген может кодировать не один белок, а целое их семейство. Ярким примером является альтернативный сплайсинг РНК гена модулятора cAMP-респонсивного элемента человека CREM, кодирующего большое семейство позитивно и негативно действующих изоформ фактора транскрипции (см. раздел 3.3.2). К аналогичным результатам приводит альтернативный процессинг полипротеинов, кодируемых некоторыми вирусными геномами. Число подобных примеров увеличивает интересный регуляторный механизм, использующий антисмысловые РНК, синтезируемые *in vivo*, в качестве регуляторных макромолекул. Антисмысловые РНК в этой роли были рассмотрены в разделе 4.2.2 в связи с регуляцией репликации плазмиды ColE1, а также как природные регуляторы трансляции мРНК у бактерий. В таких случаях обе цепи ДНК данного генетического локуса оказываются кодирующими. Следовательно, кодирование одной последовательностью нуклеотидов (одним геном в классическом понимании этого термина) множественных белков и РНК – широко распространенное явление.

Перечисленные механизмы расширяют информационную емкость генома и позволяют более компактно хранить генетическую информацию. Если в случае прокариот и, особенно, вирусов тенденция к максимальной

компактизации генома понятна и логично объясняется малыми допустимыми размерами самого генома, то этого нельзя сказать о геноме высших организмов. Действительно, неясным остается смысл эволюционного возникновения перекрывающихся генов эукариот при перенасыщении их генома некодирующими последовательностями нуклеотидов. По-видимому, это может отражать механизм образования таких генов в процессе эволюции, когда возникновение новой регуляторной последовательности (промотора, терминатора или сайта сплайсинга) внутри гена приводит к синтезу измененного, но функционально значимого полипептида. Образование новых генов по такому механизму могло бы происходить с использованием уже готовых нуклеотидных блоков существующих генов, а не из бессмысленных некодирующих последовательностей.

Запрограммированная нестабильность генов в геноме. Стабильность гена *in vivo* является одним из его жизненно важных свойств. Именно стабильность генетической информации живых организмов, проявляющаяся в сохранении фенотипических признаков в ряду поколений, делает возможным существование и самой жизни. Конечно, значение изменчивости генетического материала для эволюции в биосфере прекрасно понимали уже в самом начале развития молекулярной биологии и молекулярной генетики. Однако тогда, как и во времена де Фриза, изменение генетической информации в организме считалось редким событием, происходящим в результате генной мутации и которое чаще всего сопровождается вредными для организма последствиями, ставящими под угрозу возможность его существования.

Данные, полученные в результате работы с индивидуальными генами, обнаружили в живом организме парадоксальную ситуацию. Оказалось, что само существование большинства позвоночных животных полностью зависит не только от высокой стабильности их генома, но и от запрограммированной нестабильности ряда генетических локусов. В частности, функционирование всей иммунной системы основано на происходящих в онтогенезе животных крупномасштабных перестройках генетического материала в локусах, заключающих в себе последовательности генов иммуноглобулинов. В процессе дифференцировки В- и Т-лимфоцитов на уровне их геномной ДНК имеет место перенос последовательностей нуклеотидов (V), кодирующих переменные участки полипептидных цепей иммуноглобулинов, а также других небольших

сегментов ДНК (J и D), к последовательностям, кодирующим константные части этих белков. Объединение последовательностей нуклеотидов генов иммуноглобулинов случайным образом в разных сочетаниях, происходящее в каждой индивидуальной В- или Т-клетке по-новому, приводит к образованию огромного числа разных генов иммуноглобулинов, способных направлять биосинтез антител практически любой специфичности при распознавании антигенов. Более того, в процессе возникновения генов иммуноглобулинов *in vivo* происходят не только перестройки генома, которые сами по себе еще не обеспечивают всего разнообразия антител. При дифференцировке лимфоцитов функционирует система, осуществляющая направленное введение случайных соматических мутаций в определенные сегменты последовательностей ДНК, которые кодируют переменные участки полипептидных цепей иммуноглобулинов.

Геном клеток зародышевой линии не содержит генов, появляющихся в дифференцированных клетках иммунной системы. В онтогенезе на уровне соматических клеток происходят как бы эволюционные преобразования части генетического материала, из которой возникают новые гены и соответственно новые белки с различающимися субстратными специфичностями. Какую же часть генома в данном случае можно считать геном иммуноглобулина? Классические представления о гене не позволяют ответить на этот вопрос. Дестабилизируя соответствующие генетические локусы, организм решает задачу одновременного создания большого числа разнообразных белков с использованием потенциала генома, ограниченного в своих размерах. При таком решении организму нет необходимости из поколения в поколение осуществлять передачу громадного количества генетической информации, связанной с этими белками, биологический потенциал многих из которых не будет востребован на протяжении жизни особи следующей генерации. Потомству передается лишь генетическая программа, по которой, в принципе, могут быть синтезированы все эти белки. Таким образом, у животных в клетках зародышевого пути и в большинстве соматических клеток вообще отсутствуют гены иммуноглобулинов в классическом понимании этого термина. Они возникают *de novo* во время онтогенеза, и невозможно предсказать, какой в точности набор генов и иммуноглобулинов будет создан при реализации такой генетической программы. С помощью непредсказуемого репертуара антител

организм решает задачу защиты от не менее непредсказуемого разнообразия антигенов, которые постоянно возникают в биосфере, в том числе и вследствие природной нестабильности генов, кодирующих эти антигены. Например, запрограммированная высокая частота перестроек генов поверхностного антигена малярийного плазмодия позволяет паразиту избегать нейтрализующего влияния иммунной системы, т.е. дает возможность выжить.

В последнее время появляются данные о том, что гены, кодирующие рецепторы пахучих веществ системы обоняния позвоночных животных и человека, претерпевают перестройки, подобные таковым генов иммуноглобулинов, с теми же биохимическими последствиями. Все это указывает на то, что сейчас мы еще не осознаем до конца истинной роли индетерминизма генетической информации в жизни многоклеточных организмов как одного из основных принципов функционирования генетических систем.

В качестве еще одного примера естественной генетической нестабильности, которая приводит к созданию новых генов, не существующих у исходных организмов, рассмотрим механизм образования гена фактора σ^k (или σ^{27}) РНК-полимеразы *Bacillus subtilis*. В результате этого процесса происходит объединение в одной рамке считывания гена *spoIVCB*, кодирующего N-концевую часть σ^k , с геном *spoIIIC*, в котором закодирована недостающая C-концевая часть белка. Фактор σ^k в составе молекулы РНК-полимеразы функционирует только в процессе споруляции *B. subtilis*, обеспечивая избирательную транскрипцию бактериальных генов. При этом вновь созданный ген не передается потомству *B. subtilis*, так как материнские клетки перестают существовать после образования спор, а вместе с ними элиминируется и бактериальная хромосома. Такой способ регуляции экспрессии генов путем продуктивного объединения их участков на уровне геномной ДНК, по-видимому, распространен у клеток, претерпевающих терминальную дифференцировку. Более подробную информацию о функциональном непостоянстве генома можно найти в монографии Р.Б. Хесина.

Таким образом, генетически запрограммированное непостоянство определенных локусов во многих случаях жизненно необходимо. Только учитывая этот факт, следует рассматривать стабильность гена как важнейшее

его свойство. Кроме того, еще одним механизмом, контролирующим темп мутационных изменений отдельных генетических локусов у про- и эукариот, может быть дифференциальная защита индивидуальных генов от химического мутагенеза как следствие их специфической пространственной организации (см. раздел 5.3.3). Генетически детерминированный уровень мутагенеза отдельных локусов может определять и предпочтительное изменение генов в филогенезе. Этот же механизм мог бы контролировать мутабельность индивидуальных генов или их частей в онтогенезе, а также объяснять возможность возникновения адаптивных мутаций (см. раздел 5.1.5).

Нуклеиновые кислоты как ферменты. До появления генной инженерии РНК отводили роль промежуточных переносчиков генетической информации (мРНК, тРНК) или приписывали структурные функции (рРНК). Современная молекулярная генетика развеяла миф о том, что биологический катализ является прерогативой белковых молекул, включив в число ферментов и нуклеиновые кислоты – рибозимы и дезоксирибозимы (см. главу 9). Выясняется все большая функциональная роль в биосинтезе белка самих рибосомных РНК как катализаторов. Несмотря на это белки среди энзимов сохраняют лидирующую позицию, так как их каталитический потенциал, скрытый в гигантском разнообразии сочетаний 20 аминокислот, несопоставим с таковым ДНК, РНК и других более просто устроенных макромолекул клетки.

Современное определение гена. Несомненно, что за время активного применения методов генной инженерии представления о гене существенно изменились. Однако можно ли сказать, что понятия гена в прежнем классическом смысле этого слова больше не существует? Несмотря на соблазн согласиться с этим, приходится ответить на вопрос отрицательно. В самом деле, незыблемым остается основной принцип, заложенный в понятие гена как фрагмента нуклеиновой кислоты, в последовательности нуклеотидов которой закодирована информация о последовательности нуклеотидов в другой нуклеиновой кислоте или аминокислотной последовательности в белке. Изменение фенотипа организма однозначно связано с мутационными изменениями генотипа, т.е. изменениями последовательности нуклеотидов во фрагменте нуклеиновой кислоты (ДНК или РНК), именуемом геном. Генотипические различия, как и считалось ранее, всегда передаются от родителей потомству, т.е. носят наследуемый характер. Таким образом, на

сегодняшний день ген можно было бы определить как *наследуемую часть генома, оказывающую влияние на какой-либо фенотипический признак*. Эта формулировка по смыслу близка к классическому определению "один ген – один признак".

ЧАСТЬ II

ОСНОВНЫЕ НАПРАВЛЕНИЯ РАЗВИТИЯ ПРИКЛАДНОЙ МОЛЕКУЛЯРНОЙ ГЕНЕТИКИ

Введение

В первой части книги были кратко рассмотрены основные результаты экспериментальных и теоретических исследований по молекулярной биологии и генетике, направленные на познание молекулярных механизмов функционирования генов. Получение этих результатов стало возможным лишь благодаря разработке современных методов манипулирования генами *in vivo* и *in vitro* и, прежде всего, методов генной инженерии. Использование таких подходов, основные принципы которых будут подробнее рассмотрены ниже, привело к созданию новых фундаментальных и прикладных направлений исследований по молекулярной генетике. Реализация ряда научных программ позволила в последнее время определить полную первичную структуру генома нескольких бактерий, включая "королеву" генетики *E. coli*, *B. subtilis*, а также дрожжей *S. cerevisiae*. Полным ходом идут работы по секвенированию всего генома человека и картированию его основных известных генов. Открытие каталитической активности РНК и ДНК явилось основой нового направления в молекулярной генетике, связанного с изучением и практическим использованием рибозимов и дезоксирибозимов. Наряду с этими работами бурно развивается другой комплекс исследований, связанных с антисмысловыми РНК – мощными природными и искусственными регуляторами генной активности. Разработка многочисленных и эффективных методов направленного мутагенеза сделала возможным тонкое изучение функциональной значимости конкретных последовательностей нуклеотидов в генах и аминокислотных последовательностей в белках. Это, в свою очередь, позволило перейти к искусственному конструированию новых белков с заранее заданными свойствами, в том числе с новой ферментативной активностью. Такое направление исследований получило название *белковой инженерии*.

Введение чужеродных генов в многоклеточные организмы и осуществление их эффективной экспрессии в новом генетическом окружении привело к созданию трансгенных животных и растений. В результате был заложен прочный фундамент для *генотерапии*, а также *клонирования* целых многоклеточных организмов из отдельных соматических клеток. Современные методы молекулярной генетики революционизировали диагностику различных наследственных и приобретенных заболеваний, поскольку только они дали возможность с высокой достоверностью связать причинно-следственными отношениями конкретный генотип с конкретными патологическими состояниями организма. Такая глубина понимания молекулярных механизмов патологии недостижима другими методами и, по-видимому, лишь на этом пути возможно приближение к физиологической гармонии, т.е. к истинному здоровью человека. С учетом всего этого во второй части книги кратко рассматриваются основные достижения вышеперечисленных и некоторых других направлений молекулярной генетики, а также сделана попытка прогноза их дальнейшего развития. Изложение материала начинается с обсуждения принципов генной инженерии, так как во многом именно благодаря ей биологическая наука приобрела современный облик.

ЧАСТЬ II. ИСКУССТВЕННЫЕ ГЕНЕТИЧЕСКИЕ СИСТЕМЫ

ГЛАВА 7. ПРИНЦИПЫ ГЕННОЙ ИНЖЕНЕРИИ

Само название раздела молекулярной генетики, именуемого *генной инженерией*, указывает на то, что в результате такого рода исследований создаются искусственные генетические конструкции, в которых отдельные части генов или гены целиком объединяются в требуемой последовательности руками экспериментатора – генного инженера. Это позволяет определять их взаимное влияние и функциональное значение, а также проводить экспрессию генов в новом генетическом окружении. Таким образом, в экспериментальных условиях имеет место обмен генетической информацией как между отдельными генами организмов одного и того же вида, так и между организмами разных таксономических групп, что не происходит в природе из-за непреодолимых барьеров репродуктивной изоляции. Тем не менее, практически во всех методах современной генной инженерии фрагментарно (в адаптированном виде) используются элементы природных молекулярно-генетических механизмов.

Обмен генетической информацией между носителями этой информации – живыми организмами, а также составляющими их соматическими и половыми клетками является фундаментальным принципом существования всего живого. Половой процесс освобождает организмы от груза необратимых изменений в виде соматических мутаций и других многочисленных модификаций макромолекул, которые нарушают его нормальное функционирование в старости. Кроме того, в результате такого процесса геном нового организма воссоздается в новом сочетании аллелей, что, как правило, сопровождается расширением его адаптивных возможностей. Этот широко распространенный способ обмена генами между организмами с их передачей по вертикали, т.е. от поколения к поколению, не исчерпывает всего феномена обмена генами, непрерывно происходящего в биосфере.

Известна обширная группа генетических явлений, связанная с горизонтальной передачей генов в пределах одного поколения организмов, а также между клетками одного и того же многоклеточного организма. Характерными примерами такого рода являются специфическая и неспецифическая трансдукции, осуществляемые бактериофагами, в результате

чего происходит перенос небольших частей генома микроорганизмов. Большое значение в эволюции бактерий играет и обмен генами с помощью конъюгативных плазмид и транспозонов, в частности распространение генов устойчивости к различным химическим веществам как в популяциях родственных бактерий, так и между представителями таксономически удаленных друг от друга групп. Ретровирусы, по-видимому, и в природных условиях способны осуществлять горизонтальный перенос генов у млекопитающих, а с помощью Ti-плазмид происходит горизонтальный обмен генами и у растений. Перемещение генетической информации и изменение характера ее экспрессии возможны и в пределах самих одноклеточных и многоклеточных организмов под действием разнообразных мобильных генетических элементов, а также при воздействии мутагенных факторов окружающей среды.

Приведенные примеры показывают, что в природе обмен блоками генов, отдельными генами и их фрагментами как в пределах геномов, так и между различными геномами и организмами – обычное явление. В результате этих событий переносимые гены не только сохраняют свою способность к экспрессии в новом генетическом окружении, но и могут значительно менять ее уровень, что часто сопровождается характерным изменением фенотипа организмов.

Образование новых сочетаний генов и их частей в природных условиях, по-видимому, не носит целенаправленного характера, и лишь жесткая проверка естественным отбором может оценить жизненную значимость таких преобразований геномов. Однако сознательное использование в лабораторных условиях основных генетических принципов, лежащих в основе природных перемещений генов, позволило разработать более эффективные системы передачи генетической информации между организмами и приступить к беспрецедентным по информативности исследованиям генетических явлений на молекулярном уровне. У нас на глазах произошло рождение нового направления в молекулярной биологии – генной инженерии, значение которого не ограничивается результатами тех или иных фундаментальных и прикладных исследований. Несомненно, именно с этого момента начался новый этап эволюции биосферы Земли, все последствия которого мы в настоящее время не в состоянии предвидеть.

Необходимость манипулирования генами диктуется конкретными задачами фундаментальных и прикладных исследований. Для понимания молекулярных механизмов функционирования отдельных генов и взаимосвязанных генетических систем большое значение имеет работа с изолированными генами. Такие исследования позволяют определить границы генов, выделить их в чистом виде и идентифицировать элементы структуры, существенные для функционирования. Доказательством функциональной значимости выделенного участка генома может быть только его нормальная экспрессия в модельной генетической системе. Поэтому следующим этапом исследования выделенного гена всегда является перемещение его в такую генетическую систему, где экспрессия гена легко обнаруживается. Результаты экспрессии оценивают либо по появлению белкового продукта, кодируемого исследуемым геном, либо по изменению функций биологической системы вследствие появления в ней новой ферментативной или другой активности, например по компенсации присутствующей в этой системе мутации. Таким образом, в результате исследования структуры конкретного гена и моделирования его экспрессии в искусственной генетической системе можно понять особенности его функционирования в живом организме. Подобный подход может быть успешно применен как к известным генам, которые выделяются целенаправленно, так и к неидентифицированным ранее последовательностям нуклеотидов, функциональную значимость которых определяют лишь после выделения их в чистом виде. Последний подход реализуется в так называемой *обратной генетике*.

В настоящее время с помощью методов генной инженерии получены данные о структуре и функционировании генов разнообразных организмов, что дало возможность перейти на качественно новый уровень генетических исследований. Это, во-первых, возможность переноса гена в новое для него генетическое окружение с дальнейшей его экспрессией, что ведет к изменению свойств организма, в геном которого вводится ген (например создание продуцентов биологически активных веществ или трансгенных животных), а также осуществление генотерапии наследственных и приобретенных заболеваний путем искусственного замещения мутантных аллелей. Во-вторых, стало реальным конструирование новых генов путем объединения *in vitro* как известных, так и новых, искусственно синтезированных последовательностей

нуклеотидов. Этот подход используется в белковой инженерии для исследования функциональной значимости отдельных аминокислот и доменов в полипептидных цепях ферментов, а также для создания новых белков. В-третьих, в современной биотехнологии появилась возможность применять изолированные гены в составе генно-инженерных конструкций для получения пищевых продуктов и биологически активных веществ белковой природы.

Поскольку в экспериментальных условиях невозможно работать с одной копией гена, получение необходимого числа идентичных копий гена или его частей является первой и одной из основных задач генной инженерии. Для ее решения используют метод *молекулярного клонирования*. Сущность метода заключается в том, что нуклеотидная последовательность, которую необходимо выделить или размножить, ковалентно встраивается в самореплицирующиеся молекулы нуклеиновой кислоты, называемые *векторами*. Далее такая последовательность нуклеотидов в составе вектора вводится в клетки про- или эукариотического организма, и эти гибридные клетки *в селективных условиях*, обеспечивающих сохранение вектора внутри клеток, выращивают на питательной среде. В результате образуется клон клеток, теоретически содержащих идентичные векторные молекулы с одной и той же вставкой чужеродной последовательности нуклеотидов. Поскольку объединение молекул клонируемой последовательности нуклеотидов и вектора является не чем иным, как рекомбинацией *in vitro*, такие гибридные молекулы называют *рекомбинантными молекулами*. В настоящее время разработаны многочисленные методы, позволяющие выделять определенные последовательности нуклеотидов из сложной смеси фрагментов хромосомной ДНК, а также осуществлять обмен между строго определенными фрагментами генов и другими последовательностями нуклеиновых кислот. Во всех этих реакциях, как правило, используются высокоочищенные препараты нуклеиновых кислот и ферментов нуклеинового обмена, особенности использования которых будут кратко рассмотрены ниже.

Большинство ферментов, применяемых для молекулярного клонирования нуклеиновых кислот, участвует в метаболизме нуклеиновых кислот *in vivo*. Это означает, что генная инженерия в своем развитии опирается на достижения исследований ферментных систем метаболизма нуклеиновых кислот и существует благодаря возможности получения таких ферментов в

высокоочищенном состоянии. В то же время сам процесс клонирования и исследования клонированных последовательностей нуклеотидов сводится в основном к последовательному проведению *in vitro* определенных ферментативных реакций с использованием очищенных ферментов и их субстратов – нуклеиновых кислот.

При очистке ДНК и РНК в основном используются одни и те же приемы. Различия в методах их фракционирования определяются значительно большей чувствительностью молекул РНК к гидролитическому расщеплению из-за присутствия у рибонуклеотидов, входящих в состав РНК, свободных 2'-ОН групп, их более низкой молекулярной массой и наличием у одноцепочечных молекул характерной (компактной) пространственной структуры. Для выделения нуклеиновых кислот в нативном состоянии необходимо соблюдать, по крайней мере, две предосторожности: следует использовать мягкие условия разрушения тканей биологического объекта, содержащего изучаемые нуклеиновые кислоты, и инактивировать гидролитические ферменты (ДНКазы или РНКазы) до того, как они успеют гидролизовать молекулы нуклеиновых кислот. С учетом этого ингибиторы нуклеаз вводятся в буферные растворы до разрушения клеток или тканей. В качестве ингибиторов нуклеаз используют как неспецифические денатурирующие агенты (ионные детергенты, гидроокись ртути, гуанидинхлорид, гуанидинизотиоцианат, фенол, хлороформ и т.п.), так и специфические ингибиторы, например ингибитор РНКазы из плаценты человека или ванадиевые комплексы рибонуклеозидов.

Работу двухцепочечной геномной ДНК (особенно ДНК эукариотических клеток) сильно усложняют значительная молекулярная масса и высокая жесткость выделяемых молекул. Следствием этого являются большая вязкость растворов высокомолекулярной ДНК и легкость ее фрагментации в растворах. В связи с этим при выделении высокомолекулярной ДНК допускается лишь мягкое перемешивание ее раствора в процессе *депротеинизации* (освобождения от белков). Для этого на первых этапах очистки обычно применяют протеолитические ферменты (протеиназу К из *Trirachium album* или проназу из *Streptomyces griseus*), фенол и хлороформ. Поскольку при освобождении от белков происходит одновременная очистка ДНК и РНК, для освобождения от примесей РНК применяют высокоочищенные препараты РНКаз без примесей ДНКаз (чаще всего панкреатическую РНКазу), а в случае

выделения РНК – ДНКаз, не загрязненных РНКазами.

Разделение ДНК и РНК центрифугированием в градиенте плотности CsCl и сахарозы. Высокоэффективным методом разделения нуклеиновых кислот, широко используемым при их очистке, является центрифугирование в градиенте плотности хлористого цезия. В этом случае высокая ионная сила буферного раствора в центрифужных пробирках способствует диссоциации комплексов белков и нуклеиновых кислот, и разделение макромолекул происходит на основании их различий в плавучей плотности. Перемещение макромолекул во время центрифугирования происходит до тех пор, пока они не достигнут зоны раствора CsCl с плотностью, равной их плавучей плотности. В этой зоне происходит их концентрирование. Метод центрифугирования в градиенте плотности CsCl в присутствии бромистого этидия используется в генной инженерии для получения векторных плазмид в препаративных количествах. Он позволяет отделять суперскрученные молекулы ДНК от релаксированных и линейных, так как их плавучие плотности заметно различаются из-за разного количества молекул бромистого этидия, интеркалированного в эти топологические изомеры плазмидной ДНК. В практической генной инженерии для очистки векторных плазмид центрифугирование в градиенте плотности CsCl в настоящее время используется редко. В целях обычного применения рекомбинантных ДНК, в том числе для молекулярного клонирования, построения рестрикционных карт и определения последовательности нуклеотидов (секвенирования), разработаны методы минипрепаративного выделения. В одном из них суперскрученные молекулы рекомбинантных плазмидных ДНК легко отделяются от линейных молекул хромосомной и плазмидной ДНК в процессе щелочной денатурации с последующей быстрой нейтрализацией раствора. При этом в основном успевают ренатурировать только суперскрученные молекулы ДНК, так как две их цепи остаются физически связанными друг с другом в процессе денатурации. Образовавшийся комплекс денатурированной ДНК и белков отделяется от плазмидной ДНК центрифугированием, а от основной массы РНК освобождаются переосаждением в присутствии высокой концентрации LiCl.

Широко распространенным методом фракционирования нуклеиновых кислот по размерам молекул является центрифугирование в градиенте концентрации глицерина.

Электрофоретическое и хроматографическое разделение нуклеиновых кислот. Для дополнительной очистки плазмидной ДНК иногда применяют гель-фильтрацию на Сефакриле S1000 (крупнопористый синтетический аналог сефадекса), а также электрофорез в агарозном геле. Следует отметить, что электрофорез в агарозном, полиакриламидном или смешанном агарозно-полиакриламидном гелях используют для определения размеров и выделения небольших фрагментов двухцепочечной ДНК, образующихся под действием эндонуклеаз рестрикции или во время секвенирования ДНК. Проявление фрагментов в этом случае проводят с помощью бромистого этидия, интеркалирующего между основаниями ДНК и флуоресцирующего в ультрафиолетовом свете. Более чувствительными методами обнаружения фрагментов ДНК в гелях являются окраска серебром, авторадиграфия и введение флуоресцентной метки.

Отдельно следует упомянуть метод аффинной хроматографии на олиго(dT)-целлюлозе, используемый для специфической очистки эукариотических мРНК, большинство из которых содержит на 3'-конце поли(А)-последовательность. В олиго(dT)-целлюлозе последовательности олиго(dT)₁₂₋₁₈ ковалентно присоединены к молекулам целлюлозы своими 5'-концевыми фосфатными группами. При использовании этого метода фракцию суммарной РНК наносят на колонку с олиго(dT)-целлюлозой в условиях, при которых происходит гибридизация поли(А)-последовательностей мРНК с комплементарными им последовательностями олиго(dT)-целлюлозы. После отмывания несвязавшихся со смолой молекул РНК, большинство из которых являются рРНК и тРНК, фракцию поли(А)-РНК элюируют, прогревая колонку до температуры плавления поли(А)•олиго(dT)-гибридов или разрушая гибриды понижением ионной силы элюирующего раствора.

Все перечисленные выше способы очистки и фракционирования нуклеиновых кислот позволяют получать лишь относительно короткие последовательности нуклеотидов. Однако в настоящее время в арсенале молекулярной биологии имеются методы, которые дают возможность выделять в гомогенном состоянии последовательности нуклеотидов, эквивалентные целым хромосомам. Одним из таких методов является электрофорез в импульсном электрическом поле (pulsed field gel electrophoresis – PFGE), с помощью которого удастся препаративно разделять ДНК целых хромосом

дрожжей. Другой эффективный метод, с помощью которого можно фракционировать метафазные хромосомы млекопитающих, – это сортировка хромосом с помощью проточной цитофлуориметрии. Метод позволяет получать индивидуальные метафазные хромосомы в количестве, достаточном для конструирования клонотек генов индивидуальных хромосом.

Все упомянутые выше методы очистки нуклеиновых кислот при использовании в качестве исходного материала клеток бактерий, животных или растений дают в руки исследователей сложную смесь генов и некодирующих последовательностей нуклеотидов. Для выделения конкретных последовательностей нуклеотидов и работы с ними необходимо использовать многочисленные ферменты. Свойства некоторых ферментов, широко используемых в генной инженерии, кратко рассмотрены ниже.

7.1. Основные ферменты, используемые в генной инженерии

7.1.1. Рестриктазы и ДНК-метилазы

Среди ферментов, используемых в генной инженерии для клонирования, большое значение имеют эндонуклеазы рестрикции – *рестриктазы*. Эти ферменты, впервые открытые как часть системы рестрикции–модификации ДНК у бактерий, специфически гидролизуют молекулы двухцепочечных ДНК при наличии в них определенных последовательностей нуклеотидов, называемых *сайтами рестрикции*. В то же время метилазы используют для ограничения числа сайтов рестрикции и получения более крупных фрагментов ДНК с помощью рестриктаз.

Классификация рестриктаз. По механизму действия и молекулярной структуре различают три типа рестриктаз. Ферменты рестрикции типа I представляют собой сложные мультимерные комплексы, построенные из трех субъединиц с молекулярной массой до 300 кДа, которые обладают рестриктазной, ДНК-метилазной и АТРазной активностями. Рестриктазы типа I для проявления своей активности требуют присутствия АТФ, S-аденозилметионина и ионов Mg^{2+} , они не распознают специфические последовательности нуклеотидов и в силу этого не находят широкого применения в генной инженерии. Рестриктазы типа II узнают специфические

последовательности нуклеотидов в точке расщепления ДНК или непосредственной близости от нее, требуют для проявления активности наличия в реакционной смеси АТФ и ионов Mg^{2+} и чаще всего используются при молекулярном клонировании. Ферменты типа III также активны только в присутствии АТФ и ионов Mg^{2+} и не проявляют абсолютной зависимости от S-аденозилметионина.

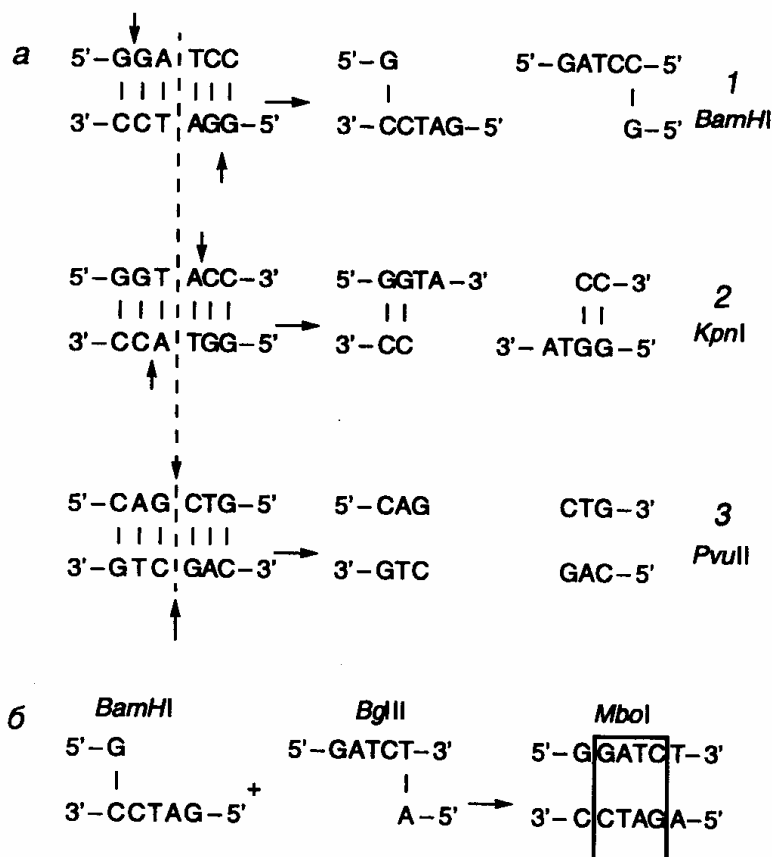


Рис. II.1. Формы разрывов двухцепочечных ДНК, образующихся под действием рестриктаз

a – 5'-выступающие (1), 3'-выступающие "липкие" (2) и "тупые" (3) концы ДНК, образующиеся под действием рестриктаз *Bam*HI, *Kpn*I и *Pvu*II соответственно. Стрелками обозначены места разрывов цепей ДНК, пунктирной линией – ось симметрии сайтов рестрикции;
б – лигирование с потерей сайта рестрикции

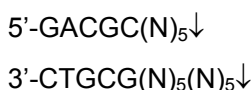
Названия рестриктаз складываются из первых букв видовых названий бактерий, в которых они обнаружены, например *Eco* – *E. coli*. В том случае,

когда различные по специфичности действия рестриктазы присутствуют в клетках разных штаммов одного вида бактерий, в название рестриктазы вводят дополнительную букву, например рестриктазы *Hinc* и *Hind* выделены из бактериальных клеток *Haemophilus influenzae*, штаммы с и d. Цифры, следующие за буквенными обозначениями, отражают последовательность открытия соответствующих рестриктаз в клетках бактерий одного вида, например *HaeI*, *HaeII* и *HaeIII* из *H. aegypticus*.

Рестриктазы типа II – основной инструмент генной инженерии. Большинство рестриктаз типа II специфически узнают на ДНК тетра- и гексануклеотидные последовательности, а по крайней мере три из них – октануклеотиды. Чем короче олигонуклеотидная последовательность сайта рестрикции, узнаваемого рестриктазой, тем чаще он встречается в случайной последовательности нуклеотидов, в которой каждый из четырех нуклеотидов представлен с одинаковой частотой (50% А–Т-пар и 50% G–C-пар). Так, случайная тетрануклеотидная последовательность встречается в среднем через каждые 256 п.о. (4^4), а гексануклеотидная – через каждые 4096 п.о. (4^6). Однако в природных ДНК распределение нуклеотидов может заметно отличаться от случайного. Например, для эукариотических ДНК характерна низкая частота встречаемости динуклеотида CpG и соответственно сайтов рестрикции, содержащих эти динуклеотиды (рестриктазы *HhaI*, *HpaII*, *TaqI*, *SalI*, *AvaI*, *HaeII*, *HindII*, *Sall*, *SmaI*, *XhoI*, *XmaI*). Существенное отклонение частоты встречаемости сайтов рестрикции от ожидаемого при случайном их распределении вдоль ДНК свойственно и хромосомам термофильных бактерий, которым, напротив, свойственно (хотя и не во всех случаях) обогащение по G–C-парам. Для большинства сайтов, узнаваемых рестриктазами типа II, характерно наличие в них симметрии второго порядка, т.е. узнаваемые ими последовательности представляют собой палиндромы, например у рестриктазы *EcoRI* – 5'-GAATTC-3'. Это означает, что нуклеотиды, расположенные в каждой из цепей на равном расстоянии от оси симметрии, комплементарны друг другу. Если точки расщепления противоположных цепей ДНК смещены друг относительно друга в сайте рестрикции, то образующиеся в результате рестрикции концы ДНК содержат выступающие одноцепочечные участки. Поскольку такие участки комплементарны сами себе и друг другу и могут между собой взаимодействовать, их часто называют "липкими" концами. В "липких"

концах выступающим одноцепочечным участком может быть как 5'-, так и 3'-конец (рис. II.1,а). Формальным признаком образования 5'- или 3'-выступающих "липких" концов в сайтах рестрикции является расположение точки расщепления цепей ДНК в последовательности, используемой для обозначения сайта рестрикции, слева или справа от оси симметрии соответственно. У некоторых рестриктаз точки расщепления обеих цепей ДНК расположены непосредственно друг под другом в сайте рестрикции. В этом случае после расщепления ДНК "липких" концов не образуется, а получаются так называемые "тупые" концы, в которых нет выступающих одноцепочечных участков ДНК (см. рис. II.1,а). Имеется одно принципиальное функциональное различие между 5'- и 3'-выступающими "липкими" концами – последние невозможно пометить путем их достройки ДНК-полимеразой. Эту особенность следует иметь в виду при выборе рестриктаз для получения рестрикционных фрагментов ДНК, которые предполагается использовать в качестве зондов.

При конструировании рекомбинантных молекул полезно помнить, что, хотя рестриктазы *Bam*HI, *Bcl*I, *Bgl*II и *Xho*II узнают разные сайты рестрикции, они образуют одни и те же "липкие" концы, GATC. То же характерно и для группы рестриктаз *Sal*GI, *Xho*I и *Ava*I (NCGA). При лигировании (см. ниже) фрагментов ДНК, образованных рестриктазами одной из таких групп, происходит их объединение, но при этом исходные сайты рестрикции теряются, так как в результате образуется новая непрерывная последовательность нуклеотидов (см. рис. II.1,б). Сайты рестрикции для некоторых рестриктаз II типа не являются симметричными. Например, рестриктаза *Hga*I узнает асимметричную последовательность 5'-GACGC-3', а одноцепочечные разрывы вносит в противоположные цепи ДНК, отступя вправо на 5 и 10 нуклеотидов соответственно:



Последовательности нуклеотидов образующихся "липких" концов являются уникальными для каждого такого сайта рестрикции. Вследствие этого рестрикционные фрагменты ДНК, образовавшиеся под действием данной рестриктазы, в смеси соединяются друг с другом лишь в строго определенной

исходной последовательности, которая задается уникальными последовательностями нуклеотидов в "липких" концах рестрикционных фрагментов ДНК. Например, при расщеплении этой рестриктазой репликативной формы ДНК фага ϕ X174 образуется 14 фрагментов, которые *in vitro* объединяются в правильную последовательность с образованием инфекционной ϕ X174-ДНК.

Универсальные рестриктазы для одноцепочечных ДНК. На основе рестриктаз, узнающих асимметричные последовательности нуклеотидов, разработана система, позволяющая расщеплять молекулы одноцепочечной ДНК в любой заданной точке. С этой целью синтезируется олигонуклеотид, 5'-концевая часть которого содержит сайт, узнаваемый такой рестриктазой, а последовательность нуклеотидов 3'-концевой части комплементарна участку ДНК, в который необходимо внести эндонуклеазный разрыв. В результате гибридизации олигонуклеотида с одноцепочечной ДНК образуется структура, изображенная на рис. II.2. При этом фермент взаимодействует с сайтом узнавания (участок I), а разрывы вносятся по местам, обозначенным стрелками. Таким образом, положение места эндонуклеазного расщепления будет целиком зависеть от последовательности нуклеотидов участка II синтетического олигонуклеотида, комплементарного ДНК-субстрату.

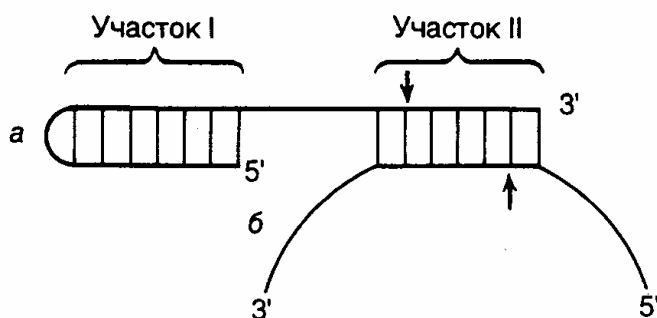


Рис. II.2. Расщепление одноцепочечной ДНК универсальной рестриктазой

а – синтетический олигонуклеотид; б – одноцепочечная ДНК.

После образования шпильки и гибридизации с одноцепочечной ДНК-мишенью олигонуклеотид образует сайт связывания рестриктазы (участок I) и сайт расщепления ДНК-ДНК-гибрида (участок II). Стрелки указывают места эндонуклеазного расщепления ДНК и олигонуклеотида

Изошизомеры. В клетках разных видов бактерий могут содержаться рестриктазы, узнающие одни и те же сайты рестрикции. Такие рестриктазы называют *изошизомерами*. Изошизомеры некоторых рестриктаз с успехом используются для обнаружения метилированных участков ДНК в геноме. Так, рестриктазы *HhaI* и *HpaII* расщепляют неметилированные последовательности GCGC и CCGG соответственно и утрачивают способность к расщеплению, если хотя бы один из остатков цитозина в этих сайтах метилирован. В то же время фермент *MspI* (изошизомер *HpaII*) расщепляет последовательность CCGG независимо от того, метилированы или неметилированы остатки цитозина в таком сайте. N-Метилирование остатков аденозина в ДНК можно обнаружить с помощью изошизомеров *Sau3A* (расщепляет как метилированные, так и неметилированные последовательности GATC), *DpnI* (расщепляет только метилированные последовательности G^{Me}ATC) и *MboI* (расщепляет только неметилированные последовательности).

Изменение специфичности действия рестриктаз в неоптимальных условиях. Рестриктазы являются высокоспецифическими ферментами. Однако для поддержания этой специфичности *in vitro* необходимо соблюдать в реакционной смеси оптимальные условия для действия ферментов. При нарушении таких условий у некоторых рестриктаз начинает проявляться вторичная (так называемая штриховая) активность. Так, рестриктаза *EcoRI* расщепляет последовательность GAATTC при pH 7,3, 100 mM NaCl в присутствии 5 mM MgCl₂, однако при изменении значений pH, понижении концентрации NaCl или замене ионов Mg²⁺ на Mn²⁺, а также в присутствии органических растворителей у фермента появляется тенденция к расщеплению более короткой последовательности AATT (так называемая активность *EcoRI*). К рестриктазам, обладающим подобными свойствами, относятся также *BamHI*, *BstI*, *BsuI*, *DdeI*, *HhaI*, *PstI*, *SalI*, *SstI*, *XbaI*.

Действие рестриктаз на необычные субстраты. Помимо двухцепочечных ДНК многие рестриктазы способны использовать ДНК-РНК-гибриды в качестве субстрата. Это относится, в частности, к рестриктазам *EcoRI*, *HindII*, *SalI*, *MspI*, *HhaI*, *AluI*, *TaqI* и *HaeIII*. Некоторые рестриктазы, например *HaeIII*, *HhaI* и *SfiI*, способны расщеплять одноцепочечную ДНК фага φX174, хотя и со значительно меньшей скоростью, чем соответствующую двухцепочечную RF-форму. Такая способность была продемонстрирована для

некоторых других рестриктаз, а также ДНК-субстратов. Остается неясным, узнают ли эти рестриктазы истинные одноцепочечные сайты или же последовательности нуклеотидов, заключенные в элементы вторичной структуры.

С развитием метода полимеразной цепной реакции (ПЦР) (см. ниже) часто возникает необходимость расщепления рестриктазами амплифицированных олигонуклеотидов недалеко от их концов, т.е. в условиях, когда сайт рестрикции фланкирован с одного из своих концов одним или несколькими нуклеотидами. В этом случае установлена четкая зависимость способности определенных рестриктаз расщеплять сайты рестрикции от количества фланкирующих сайт нуклеотидов. Данное свойство рестриктаз объясняют, в частности тем, что на самих концах двухцепочечной молекулы ДНК происходит локальное плавление двойной спирали ДНК с образованием коротких одноцепочечных участков, захватывающих сайт узнавания рестриктазами. Частично избежать локальное плавление можно понижением температуры реакционной смеси во время проведения рестрикции таких олигонуклеотидов. Поскольку эти данные имеют большое значение для практической генной инженерии, они суммированы в табл. II.1.

Таблица II.1

**Эффективность расщепления коротких последовательностей ДНК
некоторыми распространенными рестриктазами**

Рестриктаза	Последовательность олигонуклеотидов в окрестностях сайта рестрикции	Длина цепи, нт	Процент расщепления олигонуклеотида после инкубации в течение	
			2 ч	20 ч
<i>Bam</i> HI	CGGATCCG	8	10	25
	CGGGATCCCG	10	>90	>90
	CGCGGATCCCGC	12	>90	>90
<i>Bgl</i> II	CAGATCTG	8	0	0
	GAAGATCTTC	10	75	>90
	GGAAGATCTTCC	12	25	>90
<i>Cl</i> al	CATCGATT	8	0	0
	GATCGATC	8	0	0
	CCATCGATGG	10	>90	>90
	CCCATCGATGGG	12	50	50
<i>Eco</i> RI	GGAATTCC	8	>90	>90
	CGGAATTCCG	10	>90	>90
	CCGGAATTCCGG	12	>90	>90
<i>Hind</i> III	CAAGCTTG	8	0	0
	CCAAGCTTGG	10	0	0
	CCCAAGCTTGGG	12	10	75
<i>Nhe</i> I	GGCTAGCC	8	0	0
	CGGCTAGCCG	10	10	25
	CTAGCTAGCTAG	12	10	50
<i>Pst</i> I	GCTGCAGC	8	0	0
	TGCACTGCAGTGCA	14	10	10

Таблица II.1 (окончание)

Рестриктаза	Последовательность олигонуклеотидов в окрестностях сайта рестрикции	Длина цепи, нт	Процент расщепления олигонуклеотида после инкубации в течение	
			2 ч	20 ч
<i>PvuI</i>	CCGATCGG	8	0	0
	ATCGATCGAT	10	10	25
	TCGCGATCGCGA	12	0	10
<i>SacI</i>	CGAGCTCG	8	10	10
<i>SacII</i>	GCCGCGGC	8	0	0
	TCCCCGCGGGGA	12	50	>90
<i>SmaI</i>	CCCGGG	6	10	10
	CCCCGGGG	8	10	10
	CCCCCGGGGG	10	0	50
	TCCCCCGGGGGA	12	0	>90
<i>SpeI</i>	GACTAGTC	8	10	>90
	GGACTAGTCC	10	10	>90
	CGGACTAGTCCG	12	0	50
	CTAGACTAGTCTAG	14	0	50
<i>SphI</i>	GGCATGCC	8	0	0
	CATGCATGCATG	12	0	25
	ACATGCATGCATGT	14	10	50
<i>XbaI</i>	CTCTAGAG	8	0	0
	GCTCTAGAGC	10	>90	>90
	TGCTCTAGAGCA	12	75	>90
	CTAGTCTAGACTAG	14	75	>90
<i>XhoI</i>	CCTCGAGG	8	0	0
	CCCTCGAGGG	10	10	25
	CCGCTCGAGCGG	12	10	75

Примечание. Курсивом выделены последовательности сайтов рестрикции.

ДНК-метилазы. Большинство штаммов *E. coli* содержит два типа ферментов, метилирующих ДНК: *dam*- и *dcm*-метилазы. Первая осуществляет перенос метильных групп в N-положение аденина в последовательности GATC. В таком случае многие рестриктазы (например *BclI*, *MboI* или *ClaI*), в состав сайтов рестрикции которых входит данная метилированная последовательность, перестают расщеплять ДНК по этим сайтам. Аналогичное действие на некоторые рестриктазы, например *EcoRII*, оказывает и *dcm*-метилаза, осуществляющая метилирование остатков цитозина по положению C5 в последовательностях C^{Me}CAGG и C^{Me}CTGG. Для того чтобы избежать нежелательного влияния этих метилаз на клонируемые ДНК, в качестве хозяев используют мутантные штаммы *E. coli*: *dam*⁻ и *dcm*⁻. ДНК-метилазы бактериальных систем рестрикции и модификации применяют для блокирования *in vitro* соответствующих сайтов рестрикции на исследуемых фрагментах ДНК с целью получения под действием гомологичных рестриктаз фрагментов больших размеров.

7.1.2. ДНК- и РНК-лигазы

Создание фосфодиэфирных связей в одноцепочечных разрывах двухцепочечной ДНК с помощью ДНК-лигаз является наряду с рестрикцией одним из важнейших этапов получения рекомбинантных ДНК *in vitro*. Наибольшее применение в генно-инженерных исследованиях находит ДНК-лигаза бактериофага Т4. Реакция лигирования протекает в два этапа (рис. II.3). Вначале образуется промежуточный комплекс фермент–АМР (этап 1), после чего остаток АМР переносится на 5'-фосфатную группу концевого нуклеотида в точке разрыва ДНК (этап 2). Образовавшаяся фосфодиэфирная связь гидролизует во время нуклеофильной атаки 3'-ОН группы соседнего нуклеотида, что приводит к образованию новой фосфодиэфирной связи, восстанавливающей целостность сахара-фосфатного остова ДНК. Т4-ДНК-лигаза осуществляет соединение фрагментов двухцепочечной ДНК, обладающих комплементарными "липкими" или "тупыми" концами. Как следует из механизма реакции, необходимым условием протекания лигирования является наличие 5'-концевого фосфата и 3'-концевого гидроксила в точках разрыва цепей ДНК. При этом эффективность соединения фрагментов ДНК по "тупым" концам Т4-ДНК-лигазой возрастает в присутствии Т4-РНК-лигазы,

которая осуществляет ковалентное соединение 5'-фосфорилированных концов одноцепочечных ДНК или РНК с 3'-ОН группами одноцепочечных нуклеиновых кислот.

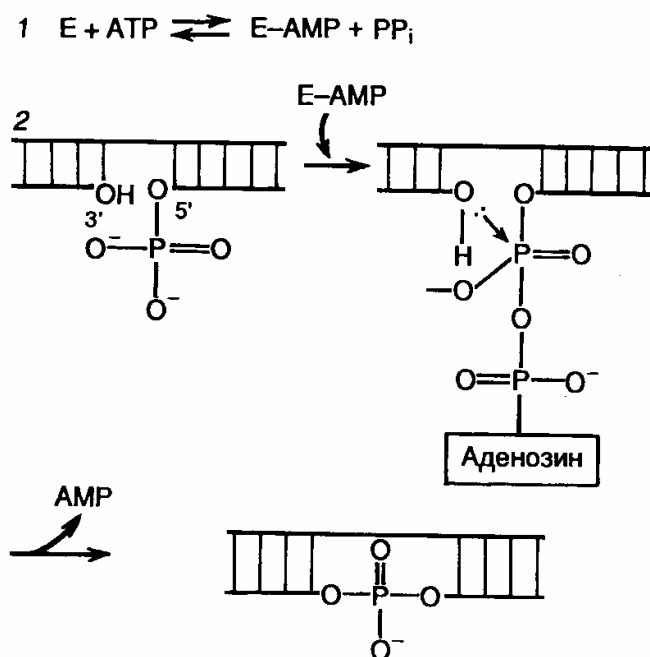


Рис. II.3. Механизм лигирования ДНК Т4-ДНК-лигазой

1 – образование промежуточного комплекса фермент Е–АМР; 2 – образование фосфодиэфирной связи

7.1.3. Ферменты матричного синтеза ДНК и РНК

К ферментам матричного синтеза нуклеиновых кислот относятся многочисленные ДНК- и РНК-зависимые ДНК- и РНК-полимеразы, осуществляющие зависимый от матричных ДНК или РНК синтез нуклеиновых кислот. Эти ферменты обычно используются в генной инженерии для получения двухцепочечных молекул ДНК из одноцепочечных, а также для обратной транскрипции, т.е. синтеза двухцепочечных ДНК, комплементарных мРНК, которые называют комплементарными ДНК (кДНК).

ДНК-зависимые ДНК-полимеразы. Среди ДНК-зависимых ДНК-полимераз наибольшее применение в генной инженерии находят ДНК-полимераза I *E. coli* и ее большой фрагмент (фрагмент Кленова), Т4-ДНК-

полимераза и в последнее время термостабильные ДНК-полимеразы, особенно ДНК-полимераза *Thermus aquaticus* (Taq-полимераза). Все эти ферменты в присутствии ионов Mg^{2+} из четырех дезоксирибонуклеозидтрифосфатов (dATP, dCTP, dGTP и TTP) осуществляют синтез ДНК, комплементарной матричной ДНК, и для функционирования требуют наличия затравки на одноцепочечной матричной ДНК, т.е. олиго- или полидезоксирибонуклеотида со свободным 3'-ОН-концом, комплементарного матричной ДНК. ДНК-полимераза I *E. coli* состоит из одной полипептидной цепи с молекулярной массой около 109 кДа и обладает тремя активностями: полимеризующей в направлении 5'→3', 5'→3'-эксонуклеазной и 3'→5'-эксонуклеазной. Большой фрагмент ДНК-полимеразы I *E. coli* (фрагмент Кленова) является частью полипептидной цепи ДНК-полимеразы I с молекулярной массой около 76 кДа, у которой отсутствует домен, соответствующий 5'→3'-эксонуклеазе. Как ДНК-полимераза I, так и ее фрагмент используются для введения радиоактивно меченных дезоксирибонуклеотидов в синтезируемые цепи ДНК путем *ник-трансляции*, т.е. перемещения одноцепочечного разрыва вдоль молекулы двухцепочечной ДНК, в котором 3'-ОН-конец используется в качестве затравки для ферментов. При этом ДНК-полимераза I прокладывает себе путь с помощью 5'→3'-эксонуклеазы, а фрагмент Кленова вытесняет цепь ДНК с 5'-конца. Кроме того, фрагмент Кленова используют для синтеза второй цепи кДНК, секвенирования ДНК по методу Сенгера, заполнения 5'-выступающих "липких" концов ДНК с образованием "тупых" концов, введения концевой радиоактивной метки, а также для удаления 3'-выступающих концов рестрикционных фрагментов ДНК 3'→5'-эксонуклеазой этого фермента. Как и ДНК-полимераза I *E. coli*, T4-ДНК-полимераза обладает 3'→5'- (но не 5'→3'-) эксонуклеазной активностью, которая у последней, по крайней мере, в 200 раз выше. Это позволяет использовать T4-ДНК-полимеразу, в частности, для введения радиоактивной метки путем реакции обмена немеченного 3'-концевого нуклеотида на меченный радиоактивным изотопом.

Полимеразная цепная реакция (ПЦР). Термостабильная Taq-полимераза в настоящее время широко используется для проведения *полимеразной цепной реакции (ПЦР)*, а ее модифицированный аналог – и для секвенирования ДНК по методу Сенгера. Сущность ПЦР заключается в

следующем (рис. II.4). Реакционная смесь обычно содержит исследуемую ДНК, dATP, dCTP, dGTP и TTP, Taq-полимеразу и два синтетических олигодезоксирибонуклеотидных праймера длиной 15–30 нуклеотидов, комплементарных участкам противоположных цепей ДНК, фланкирующих исследуемый участок ДНК-матрицы. ПЦР начинается с кратковременного (1–2 мин) прогрева реакционной смеси при $\sim 95^{\circ}$, в процессе которого происходит плавление цепей матричной ДНК и инактивация примесных белков, которые могут присутствовать в реакционной смеси, если используется недостаточно очищенная матричная ДНК, например из клинических образцов. Далее реакционную смесь охлаждают до температуры отжига праймеров с матричной ДНК ($37\text{--}65^{\circ}$). В результате праймеры связываются с комплементарными им местами на матрице и с этого момента начинают служить затравками для синтеза Taq-полимеразой новых цепей ДНК, комплементарных каждой из цепей денатурированной матричной ДНК. Для проведения синтеза ДНК в оптимальных для Taq-полимеразы условиях температуру реакционной смеси поднимают до 72° и при такой температуре проводят элонгацию цепей вновь синтезирующейся ДНК в течение 0,5–2 мин. По окончании стадии синтеза ДНК заканчивается первый цикл ПЦР, после чего проводятся такие же второй, третий и т.д. в автоматическом режиме. По окончании третьего цикла уже образуется дискретный продукт ПЦР – фрагмент двухцепочечной ДНК, содержащий на своих 5'-концах последовательности нуклеотидов праймеров. Количество продукта ПЦР в реакционной смеси после завершения каждого цикла удваивается и, следовательно, по мере прохождения реакции экспоненциально возрастает, достигая нескольких микрограммов в реакционной смеси объемом 25–50 мкл. По окончании ПЦР содержимое реакционной смеси анализируют электрофорезом в агарозном геле в присутствии бромистого этидия, где продукт реакции выявляется в УФ-свете. Продукт ПЦР можно выделить из агарозного геля в чистом виде и работать с ним, как с обычным фрагментом двухцепочечной ДНК, т.е. гидролизовать рестриктазами, клонировать в плазмидных векторах по "липким" и "тупым" концам, секвенировать или использовать в качестве зондов при гибридизации.

С помощью ПЦР можно амплифицировать *in vitro* сегменты ДНК длиной от 0,1 до 5–7 т.п.о. и более, а для получения положительного результата достаточно присутствия в реакционной смеси одной–двух копий

амплифицируемой последовательности нуклеотидов (например геномной ДНК, содержащейся в одной–двух соматических клетках). При этом не возникает необходимости в тщательной очистке матрицы, так как большинство попадающих в реакцию смесь белков и ферментов инактивируется в первых же циклах ПЦР и не оказывает влияния на протекание реакции при высоких температурах. Благодаря таким особенностям, как простота постановки, высокая чувствительность и воспроизводимость, ПЦР в последнее время получила широкое распространение в фундаментальных и прикладных исследованиях по молекулярной биологии и генетике. Возможность амплификации любого сегмента ДНК, последовательность нуклеотидов которого известна, и получение его по завершении ПЦР в гомогенном виде и препаративном количестве делают ПЦР альтернативным методом молекулярного клонирования коротких фрагментов ДНК. При этом не возникает необходимости в применении тех сложных методических приемов, которые используют в генной инженерии при обычном клонировании.

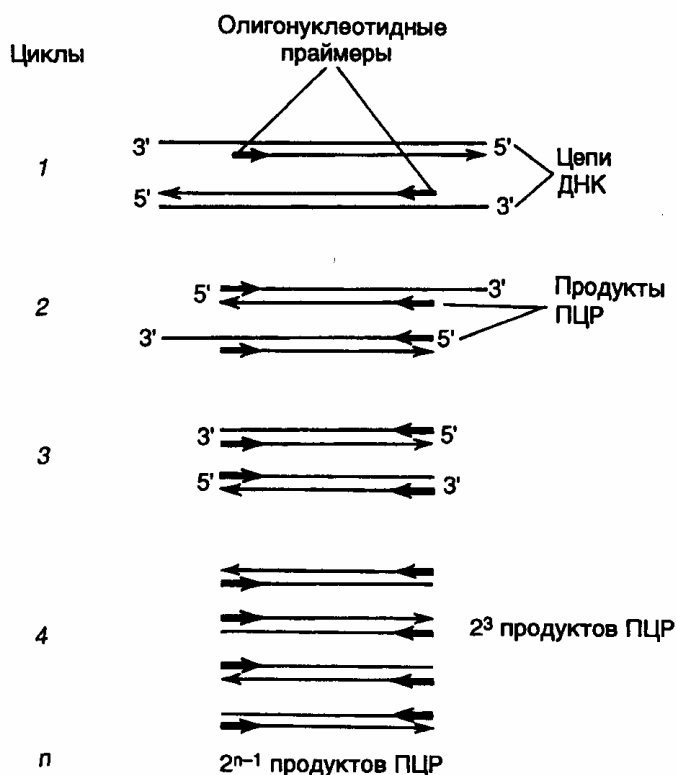


Рис. II.4. Полимеразная цепная реакция (ПЦР)

Поскольку специфичность ПЦР зависит главным образом от соответствия последовательности нуклеотидов праймеров первичной структуре амплифицируемых фрагментов ДНК, эта специфичность приближается к абсолютной при повышении длины праймеров до 17–20 нуклеотидов. Такая высокая специфичность и чувствительность позволяют использовать ПЦР для ДНК-диагностики инфекционных и наследственных заболеваний человека, а также вирусоносительства в тех случаях, когда вирусом в латентной форме заражены лишь отдельные соматические клетки, например в случае вируса иммунодефицита человека (ВИЧ) (см. главу 11). Разработка метода ПЦР во многом расширила методические возможности молекулярной генетики, и в частности генной инженерии, причем настолько, что это кардинально изменило и усилило исследовательский потенциал многих ее направлений.

Точность амплификации ДНК термостабильными ДНК-полимеразами является особенно критическим параметром в том случае, если продукты ПЦР используются далее для конструирования экспрессирующихся генов. Необходимость в уменьшении частоты мутаций стимулирует поиски новых термостабильных ДНК-полимераз, амплифицирующих ДНК с большей точностью. В табл. II.2 приведены результаты испытаний точности синтеза ДНК различными коммерческими ДНК-полимеразами.

Относительно высокая точность синтеза ДНК *in vitro* ДНК-полимеразами *Pfu* и *Vent* обусловлена наличием у них 3'-5'-экзонуклеазной корректирующей активности. Частота ошибочно включенных нуклеотидов в результате одновременного функционирования двух разных ДНК-полимераз в смесях, применяемых для амплификации длинных сегментов ДНК (long PCR), оказывается ниже, чем при использовании Taq-полимеразы, но выше, чем для очищенной ДНК-полимеразы *Pfu*.

При проведении аллель-специфической ПЦР (см. раздел 11.1) необходимо учитывать способность Taq-полимеразы инициировать синтез ДНК на праймерах, 3'-концевой нуклеотид которых некомплементарен матричной ДНК. При исследовании этого вопроса оказалось, что не все пары

Таблица II.2

**Частота ошибок при синтезе ДНК, осуществляемом
термостабильными ДНК-полимеразами in vitro при проведении
ПЦР в оптимальных условиях**

ДНК-полимераза	Частота мутаций (на 1 нуклеотид/1 раунд репликации)
Pfu	$1,3 \cdot 10^{-6}$
Deep Vent	$2,7 \cdot 10^{-6}$
Vent	$2,8 \cdot 10^{-6}$
Taq	$8,0 \cdot 10^{-6}$
exo(-)Pfu и UITma	$5,0 \cdot 10^{-5}$
Смесь Taq/Pfu или KlenTaq/Pfu	$\sim 5 \cdot 10^{-6}$

ошибочно спаренных нуклеотидов праймера и ДНК-матрицы одинаково эффективно предотвращают синтез ДНК *Taq*-полимеразой. По способности направлять синтез ДНК все такие пары 3'-концевого нуклеотида праймера и соответствующего нуклеотида матричной ДНК располагаются в следующий ряд: C–C, A–G, G–A, G–G < A–A < T–C, C–T, T–T << A–C, C–A << G–T, T–G (данные фирмы "Cetus"). При этом наименее благоприятными для продолжения синтеза ДНК являются пары нуклеотидов пурин–пурин и пиримидин–пиримидин.

РНК-зависимые ДНК-полимеразы (обратные транскриптазы, ревертазы). Обратные транскриптазы способны осуществлять синтез ДНК на матрице РНК, полимеризуя четыре дезоксирибонуклеозидтрифосфата, как это имеет место в случае ДНК-зависимых ДНК-полимераз. Обратные транскриптазы, так же как и ДНК-полимеразы функционируют только при наличии затравки. Широко используемая в генной инженерии обратная транскриптаза вируса миелобластома птиц помимо полимеризующей активности обладает 5'→3'- и 3'→5'-экзорибонуклеазными активностями и расщепляет РНК в ДНК-РНК-гибридах по процессивному механизму. Обратные транскриптазы находят применение в синтезе двухцепочечных ДНК, комплементарных РНК (особенно мРНК), для последующего ее клонирования в плазмидных векторах при получении библиотек (клонотек) кДНК (см. раздел 7.3). Обратные транскриптазы, подобно ДНК-полимеразам, могут быть использованы для введения радиоактивной или флуоресцентной метки в ДНК-зонды в составе соответствующим образом меченных дезоксирибонуклеозидтрифосфатов.

Недавно способность синтезировать ДНК на матрице РНК в определенных условиях была продемонстрирована для термостабильной ДНК-полимеразы *Thermus thermophilus*. Это позволяет использовать ее для прямого обнаружения специфических РНК в биологических образцах методом ПЦР. Современные модификации такого подхода дают возможность в одной реакционной смеси (и пробирке) синтезировать в реакции обратной транскрипции небольшое число копий амплифицируемого фрагмента ДНК на матрице РНК, которые сразу же используются тем же ферментом в качестве матрицы в обычной ПЦР (one tube PCR).

7.1.4. Другие ферменты

Среди других многочисленных ферментов, используемых в генной инженерии, прежде всего следует упомянуть **полинуклеотидкиназы**, которые осуществляют перенос γ -фосфатных групп АТР на 5'-ОН группы ДНК или РНК. *Полинуклеотидкиназа бактериофага T4* широко используется для введения радиоактивной метки в ДНК или РНК с целью получения радиоактивно меченных зондов или секвенирования нуклеиновых кислот.

Терминальная трансфераза из тимуса телят (терминальная дезоксирибонуклеотидилтрансфераза), осуществляющая последовательное присоединение дезоксирибонуклеозидмонофосфатов из пула дезоксирибонуклеозидтрифосфатов к 3'-ОН-группам молекул ДНК, используется для введения радиоактивной метки в составе меченых нуклеотидов в 3'-концы ДНК, а также присоединения к 3'-концам фрагментов ДНК (особенно кДНК) протяженных гомополимерных последовательностей нуклеотидов (*коннекторов*) для последующего их клонирования.

Щелочные фосфатазы бактерий (BAP) и кишечника телят (CIAP) катализируют удаление 5'-фосфатных групп ДНК или РНК, а также расщепление макроэргических связей рибо- и дезоксирибонуклеозидтрифосфатов. Их используют при подготовке фрагментов нуклеиновых кислот к введению 5'-концевой радиоактивной метки ^{32}P , а также для предотвращения лигирования векторных молекул ДНК самих на себя. Как уже упоминалось выше, ДНК-лигаза способна образовывать фосфодиэфирные связи в одноцепочечных разрывах лишь при наличии в них 5'-концевого фосфата. Удаление 5'-концевых фосфатных групп молекул вектора, лианеризованных с помощью одной рестриктазы во время его подготовки для клонирования соответствующего фрагмента ДНК, предотвращает образование кольцевых молекул вектора (без вставки), а также его олигомеров во время лигирования с клонируемой последовательностью. Необходимые для лигирования со вставкой фосфатные группы содержатся в самих клонируемых фрагментах ДНК, а остающиеся в результате неполного лигирования два одноцепочечных разрыва репарируются *in vivo* после введения вектора со вставкой в бактериальные клетки.

В заключение следует рассмотреть некоторые **нуклеазы**, часто

используемые в генной инженерии. Нуклеазы – это гидролитические ферменты, расщепляющие нуклеиновые кислоты. По механизму расщепления субстратов нуклеазы разделяют на экзо- и эндонуклеазы. *Эксонуклеазы* осуществляют последовательное отщепление моно- или небольших олигонуклеотидов с концов молекул ДНК или РНК, тогда как *эндонуклеазы* вносят в молекулы нуклеиновых кислот внутренние разрывы. В частности, к эндонуклеазам относятся многочисленные рестриктазы, рассмотренные выше. Некоторые нуклеазы обладают более широкой специфичностью и могут гидролизовать как РНК, так и ДНК, а также одновременно проявлять эндо- и экзонуклеазную активности. Нуклеаза Bal31 из *Alteromonas asperjiana* последовательно отщепляет отдельные моно- и олигонуклеотиды с 5'- и 3'-концов двухцепочечных ДНК, укорачивая обе цепи ДНК приблизительно с равной скоростью. При этом нуклеаза Bal31 гидролизует и одноцепочечные ДНК.

Эксонуклеаза III *E. coli* катализирует последовательное отщепление 5'-нуклеотидов из двухцепочечных ДНК в направлении 3'→5'. Фермент обладает эндонуклеазной активностью по отношению к апуринизированной ДНК, активностью РНКазы H (гидролиз РНК в РНК-ДНК-гибридах) и 3'-фосфатазной активностью.

Эксонуклеаза фага λ катализирует последовательное отщепление 5'-моонуклеотидов в двухцепочечных ДНК при наличии в них 5'-концевых фосфатных групп. Ее используют для получения одноцепочечных молекул ДНК с целью их последующего секвенирования.

Нуклеаза S1 из *Aspergillus orizae* специфически расщепляет одноцепочечные молекулы ДНК с образованием 5'-фосфорилированных моно- и олигонуклеотидов. Те же свойства присущи и нуклеазе золотистой фасоли (mung bean).

Панкреатическая рибонуклеаза А (РНКазы А) обладает активностью эндорибонуклеазы, специфически расщепляющей фосфодиэфирные связи, образованные пиримидиновыми нуклеотидами. Продуктами гидролиза являются 3'-фосфорилированные пиримидиновые моонуклеотиды и олигонуклеотиды, содержащие концевые пиримидин-3'-монофосфаты.

Панкреатическая дезоксирибонуклеаза I (ДНКазы I) представляет собой эндонуклеазу, гидролизующую как одно-, так и двухцепочечную ДНК с образованием сложной смеси моно- и олигонуклеотидов, содержащих 5'-

фосфатные группы. В присутствии ионов Mg^{2+} ДНКазы I независимо атакует каждую цепь ДНК, при этом места разрывов располагаются статистически вдоль молекулы, а в присутствии ионов Mn^{2+} расщепляет обе цепи ДНК приблизительно напротив друг друга.

Рассмотренные ферменты не исчерпывают всего списка белков, используемых в генной инженерии. Особенности применения многих из этих ферментов для решения конкретных генно-инженерных задач будут не раз обсуждаться при дальнейшем изложении принципов генной инженерии.

7.2. Векторы

Ферменты, описанные в предыдущем разделе, позволяют производить тонкие манипуляции как с протяженными молекулами ДНК, так и с их фрагментами. В частности, с помощью рестриктаз можно с большой точностью разрезать молекулы ДНК, а образовавшиеся в результате фрагменты соединять в любой желаемой последовательности друг с другом, восстанавливая сахаро-фосфатный остов молекулы ДНК с помощью ДНК-лигазы. Однако с использованием только этих ферментов еще нельзя решить одну из основных методических задач молекулярной генетики – выделение любой требуемой нуклеотидной последовательности в чистом виде и в количестве, достаточном для исследования этих последовательностей биохимическими методами. Исключение составляет метод ПЦР, однако его применение ограничивается короткими последовательностями нуклеотидов.

Основная идея, позволяющая решать эту задачу, заключается в том, чтобы присоединить исследуемые фрагменты ДНК к молекуле-переносчику, которая могла бы автономно существовать внутри бактериальных или эукариотических клеток в виде одной или нескольких копий и передаваться вместе со встроенным в нее фрагментом ДНК от одной клетки к другой. Такие молекулы-переносчики фрагментов нуклеиновых кислот были созданы, их называют *векторами*, и они являются одним из важнейших инструментов генной инженерии.

Идеальная векторная молекула должна обладать несколькими обязательными свойствами. Во-первых, любой вектор должен длительное время существовать в популяции клеток-хозяев, т.е. реплицироваться автономно или вместе с хромосомами клеток. Во-вторых, в любом векторе

должны быть биохимические или генетические маркеры, которые позволяли бы обнаруживать его присутствие в клетках. В-третьих, структура векторной молекулы должна допускать встраивание в нее чужеродной последовательности нуклеотидов без нарушения ее функциональной целостности. Для конструирования векторов в генной инженерии используют небольшие молекулы нуклеиновых кислот, способные к автономной репликации в бактериальных и эукариотических клетках – плазмиды, хромосомы вирусов, а также фрагменты хромосом эукариотических клеток.

7.2.1. Плазмидные векторы

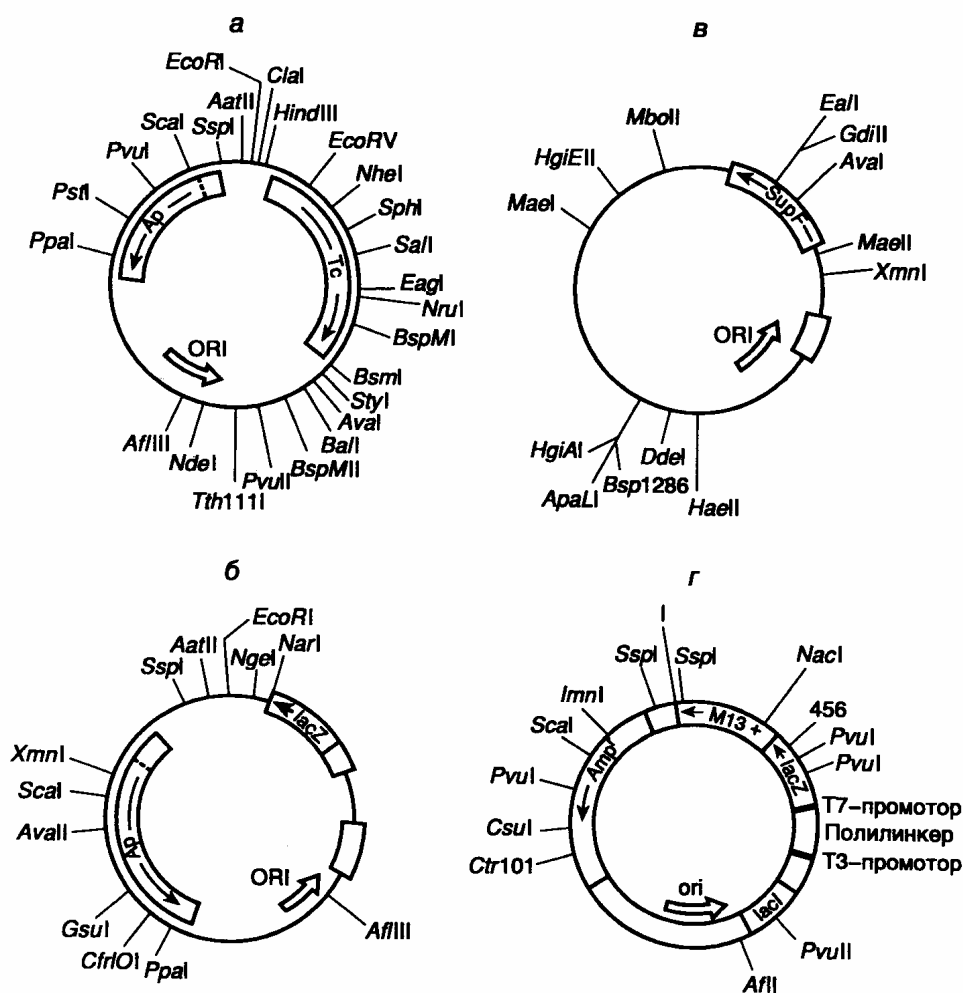


Рис. II.5. Различные векторы для клонирования ДНК и их рестрикционные карты

Обозначены положения уникальных сайтов рестрикции, а также функционально значимых генов

а – векторная плазмида pBR322; *б* – экспрессирующая векторная плазмида pUC18; *в* – векторная плазмида π AN7, предназначенная для отбора рекомбинантных клонов с использованием гомологичной рекомбинации; *г* – многофункциональная векторная плазмида Bluescript

Первые эффективные векторы для клонирования фрагментов чужеродной ДНК, не утратившие своего значения и поныне, были получены с использованием бактериальных плазмид. Большая серия векторных плазмид, обозначенных символом pBR, создана на основе репликона природной плазмиды ColEI, придающей клеткам *E. coli* устойчивость к колицину путем его объединения с генами устойчивости к антибиотикам. Таким образом, бактериальные клетки, несущие подобные комбинированные плазмиды, приобретали устойчивость к соответствующим антибиотикам, и их было легко отличить от бесплазмидных клеток путем простого посева на питательную среду с антибиотиками. Генетическая карта одного широко распространенного вектора этой серии – pBR322 изображена на рис. II.5,а. Такая плазмида представляет собой кольцевую ковалентно замкнутую молекулу ДНК длиной 4363 п.о. Последовательность нуклеотидов pBR322 полностью известна. Плазмида содержит гены устойчивости к тетрациклину (Tc) и ампициллину (Ap), которые были перенесены в плазмиду pBR322 из плазмиды pSC101 и транспозона Tn3 соответственно. Оба этих гена являются селективируемыми генетическими маркерами плазмиды, т.е. позволяют проводить отбор бактериальных клеток с плазмидой pBR322 по их способности к росту на питательных средах в присутствии тетрациклина и(или) ампициллина. Плазмида pBR322 содержит также обеспечивающий ее стабильную репликацию в клетках *E. coli* участок ДНК, который включает область начала репликации (*ori*). Характерной чертой плазмиды pBR322, как и любого современного вектора, является наличие в ней нескольких уникальных сайтов рестрикции, обозначенных на генетической карте. Следует иметь в виду, что встраивание в плазмиду клонируемых чужеродных фрагментов ДНК по сайтам рестрикции, расположенным в генах Ap или Tc (например *Pst*I или *Bam*HI), будет нарушать целостность этих генов и их функциональную активность. В результате происходит утрата бактериальными клетками, содержащими рекомбинантные плазмиды, устойчивости к соответствующим антибиотикам. По

такому признаку легко различить бактериальные клетки, не содержащие плазмиды (не растут в присутствии ампициллина и тетрациклина), клетки с плазмидой, не содержащей вставки клонируемой ДНК (растут в присутствии обоих антибиотиков), и клетки с рекомбинантными плазмидами (в зависимости от локализации вставки могут расти на среде только с одним из двух вышеупомянутых антибиотиков). Следовательно, наличие в векторных молекулах селективируемых маркеров резко повышает эффективность клонирования из-за возможности проведения быстрого отбора рекомбинантных плазмид на селективных питательных средах.

Помимо генов устойчивости к антибиотикам в качестве селективируемых маркеров используют и другие гены или их фрагменты. В частности, для этих целей часто применяются гены различных ферментов, присутствие которых в клетках в составе плазмиды обнаруживают по появлению соответствующей ферментативной активности. В часто используемых векторах серии pUC таким селективируемым маркером является 5'-концевая часть гена β -галактозидазы *E. coli* – *lacZ'* (см. рис. II.5,б). Эта часть гена, находящаяся под контролем *lac*-промотора, кодирует N-концевую часть β -галактозидазы (так называемый α -пептид), которая путем объединения с недостающей C-концевой частью полипептида без образования пептидной связи восстанавливает ферментативную активность β -галактозидазы. β -Галактозидаза обладает способностью расщеплять искусственный субстрат Xgal (5-бром-4-хлор-3-индолил- β -D-галактопиранозид) с образованием окрашенного в голубой цвет продукта реакции. В том случае, когда в бактериальных клетках, которые содержат в хромосоме недостающую экспрессирующуюся 3'-концевую часть гена *lacZ* и выращиваются на среде с Xgal, в результате комплементации α -пептидом вектора появляется активность β -галактозидазы, образованные этими клетками бактериальные колонии окрашиваются в голубой цвет. Уникальные сайты рестрикции для клонирования ДНК локализованы в начале гена β -галактозидазы в составе *полилинкера* (синтетической последовательности нуклеотидов, содержащей несколько перекрывающихся уникальных сайтов рестрикции), который не нарушает функциональной целостности последовательности нуклеотидов α -пептида. Вставка рекомбинантной ДНК в эти векторы разрывает структурную часть гена β -галактозидазы и инактивирует

его, в связи с чем колонии бактерий с подобными рекомбинантными плазмидами, выросшие на питательной среде с Xgal, не окрашены. Поскольку векторы серии pUC одновременно содержат и ген устойчивости к ампициллину, отбор бактерий, несущих рекомбинантные плазмиды, можно проводить одновременно по этим двум маркерам. На питательной среде с ампициллином и Xgal вырастают только бактерии, устойчивые к антибиотику, т.е. содержащие плазмиду pUC, а среди выросших колоний лишь неокрашенные содержат вставку чужеродной ДНК. В качестве селективируемых маркеров в векторных молекулах часто используются гены, присутствие которых может быть обнаружено косвенно по комплементации генетических дефектов бактериальных клеток-хозяев, что делает их жизнеспособными в определенных селективных условиях. Подробнее об одной из таких систем см. в разделе 7.3.3.

За короткий период развития генной инженерии было сконструировано труднообозримое количество векторных плазмид, обеспечивающих конкретные потребности исследователей. Одной из вершин генно-инженерного искусства, прекрасно иллюстрирующей возможности генной инженерии, в настоящее время являются полифункциональные векторы серии Bluescript, полученные фирмой "Stratagene" (США) (см. рис. II.5,а). Вектор Bluescript M13+ представляет собой кольцевую ковалентно замкнутую молекулу ДНК длиной около 3 т.п.о. Он включает в себя ген устойчивости к ампициллину Amp^r , ген β -галактозидазы *lacZ*, в N-концевую часть которого встроен полилинкер, содержащий уникальные сайты рестрикции для 21 рестриктазы, промоторно-операторную область *lacZ*, а также ген *lac*-репрессора *lacI*. В результате встраивания клонируемого фрагмента ДНК в полилинкер происходят разрыв кодирующей части гена *lacZ* и инактивация β -галактозидазы, что, как и в случае вектора pUC18, можно обнаружить по исчезновению окраски колоний бактерий, содержащих этот вектор со вставкой клонированной ДНК. Кроме того, встроенный в полилинкер фрагмент ДНК попадает под контроль промоторно-операторной регуляторной последовательности гена *lacZ* и в присутствии индуктора IPTG может быть экспрессирован в клетках *E. coli*. В дополнение к этому полилинкер в векторной плазмиде содержит на одном конце промотор для T7-, а на другом – для T3-РНК-полимераз, которые ориентированы навстречу друг другу. Это позволяет транскрибировать любую из цепей

клонированного фрагмента ДНК *in vitro* с помощью той или другой РНК-полимеразы и получать препаративные количества мРНК или же комплементарной ей антисмысловой РНК. Кроме того, вектор Bluescript M13+ обладает межгенной областью (IG) фага f1, родственного фагу M13. Эта область детерминирует все *цис*-действующие функциональные последовательности нуклеотидов фага, необходимые для репликации его хромосомы и упаковки ее в фаговые частицы. В присутствии фага-помощника M13 происходит преимущественная упаковка образовавшейся в результате репликации одноцепочечной плазмиды в фаговые частицы M13. Одноцепочечная ДНК Bluescript M13+ после очистки может быть использована непосредственно для секвенирования клонированной ДНК или проведения сайт-специфического мутагенеза. Векторы типа Bluescript M13+, способные существовать либо в виде плазмиды, либо в составе фаговых частиц *нитевидных бактериофагов*, называют *фагмидами*.

7.2.2. Векторы на основе фага λ

Основным недостатком плазмидных векторов для клонирования является их малая емкость в отношении клонируемых фрагментов ДНК. Размер вставок клонируемой ДНК в плазмидных векторах, которые способны стабильно в них существовать, как правило, не превышает нескольких тысяч пар оснований. Большие вставки ДНК в векторных плаزمиде неустойчивы, и их размеры постепенно уменьшаются по мере увеличения числа раундов репликации таких рекомбинантных плазмид *in vivo*. Преимущественное делетирование чужеродной ДНК в плазмиде большого размера связано с тем, что в бактериальных клетках селективное преимущество получают те плазмиды, время репликации которых минимально. Поэтому нуклеотидные последовательности ДНК, не участвующие в репликации векторных плазмид, постепенно элиминируются посредством делеций при длительном культивировании рекомбинантных бактерий.

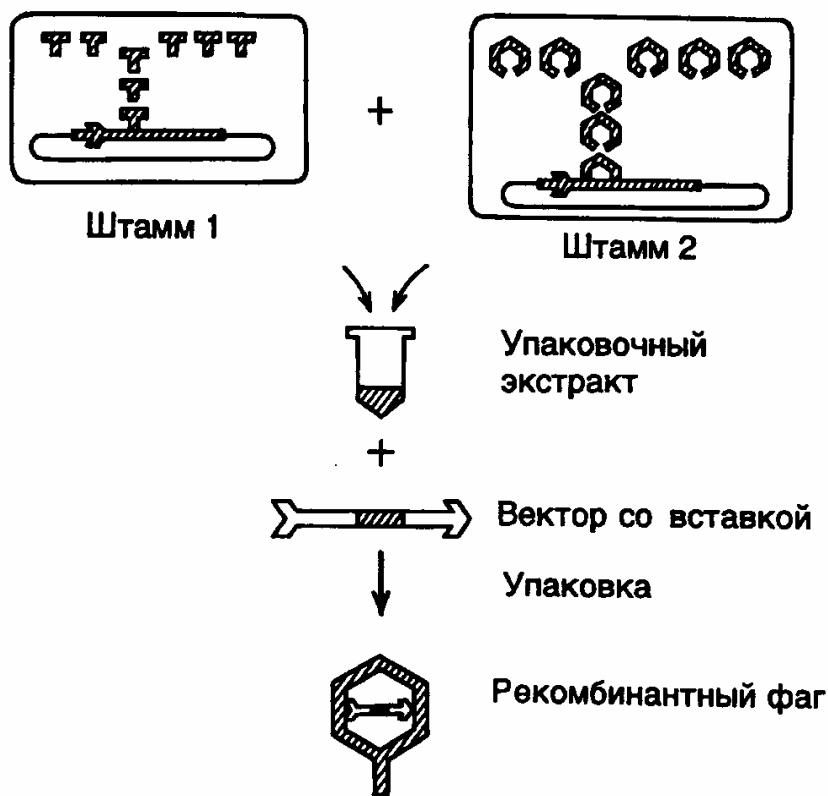


Рис. II.6. Упаковка рекомбинантной фаговой ДНК в фаговые частицы *in vitro*

Емкость клонирующих векторов была значительно повышена с появлением векторов, сконструированных на основе хромосомы бактериофага λ . Получившие широкое распространение векторы серий Charon, λ gt11 и EMBL обладают, по крайней мере, двумя существенными преимуществами перед плазмидными векторами. Во-первых, векторы на основе ДНК фага λ обладают значительно большей емкостью, в них можно клонировать фрагменты ДНК длиной от 5 до 25 т.п.о. Во-вторых, фаговые частицы, содержащие упакованную ДНК, способны проходить литический цикл развития внутри бактериальных клеток и, следовательно, образовывать стерильные пятна (бляшки) на газоне бактерий. Такие бляшки содержат в концентрированном виде как сами фаговые частицы с упакованными в них рекомбинантными молекулами ДНК, так и все продукты метаболизма зараженных бактериальных клеток, включая белки и ферменты, которые появляются в результате экспрессии клонированных

бактериальных генов. Каждая бляшка возникает вследствие развития индивидуальной фаговой частицы, содержащей рекомбинантную ДНК только одного типа, а, следовательно, все фаговые частицы одной бляшки ($\sim 10^{10}$) представляют собой, как правило, клон идентичных фаговых частиц (они могут различаться в редких случаях за счет мутационных изменений их генома, произошедших в процессе жизненного цикла фага, либо в том случае, если одна бактериальная клетка заражается несколькими фаговыми частицами одновременно). Все это позволяет легко обнаруживать в фаговых бляшках искомые ферментативные активности или последовательности нуклеотидов и идентифицировать клонированные последовательности ДНК. В основе конструирования фаговых векторов лежит несколько простых принципов (рис. II.6). В середине молекулы λ -ДНК длиной ~ 45 т.п.о. расположен участок хромосомы (~ 15 т.п.о.), который не является необходимым для литического развития бактериофага. Поэтому, в принципе, его можно заменить на любой фрагмент ДНК аналогичного размера и осуществить клонирование фрагмента путем размножения рекомбинантного бактериофага. Поскольку механизм упаковки хромосомной ДНК в фаговые частицы основан на включении ДНК строго определенного размера, рекомбинантные ДНК, содержащие фрагменты клонируемой ДНК, которые не соответствуют оптимальному размеру, не упаковываются и не клонируются. Это позволяет легко освобождаться от фаговых частиц, не содержащих вставки клонируемой ДНК, и оптимизировать процесс клонирования путем снижения в упаковочных экстрактах доли нежизнеспособных фаговых частиц. Процесс упаковки фаговой ДНК в зрелые фаговые частицы осуществляется в смеси бесклеточных экстрактов двух штаммов *E. coli*, лизогенных по дефектным бактериофагам λ . В одном штамме амбер-мутацией инактивирован один из белков фагового капсида (продукт гена *E*), а в другом – ген *A*, продукт которого необходим для включения фаговой ДНК в головку бактериофага. Имеются и другие пары лизогенных штаммов *E. coli*, позволяющие производить упаковку ДНК в фаговые частицы с использованием тех же общих принципов. Объединение бесклеточных лизатов обоих штаммов *E. coli* приводит к взаимной комплементации недостающих функций с помощью соответствующих белков дикого типа. Таким образом, в объединенных экстрактах имеются все компоненты, необходимые для сборки зрелых

инфекционных фаговых частиц, в них происходит упаковка рекомбинантной ДНК с эффективностью образования 10^4 – 10^5 фаговых частиц на 1 мкг упаковываемой ДНК. Помимо вышеупомянутых мутаций ДНК λ -лизогенов содержат температурно-чувствительную мутацию в репрессоре *ci*, который инактивируется после переноса лизогенных клеток *E. coli* на непермиссивную температуру (42°), что сопровождается индукцией профага λ и накоплением внутри бактериальных клеток белковых продуктов, необходимых для упаковки ДНК. ДНК профагов также содержит делецию *b2*, элиминирующую сайт *att*, необходимый для интеграции фаговой ДНК в бактериальную хромосому. Это предотвращает выход ДНК профага из бактериальной хромосомы, а следовательно, и ее упаковку *in vitro*. Кроме того, в хромосоме профага имеется мутация, инактивирующая ген *S*, кодирующий лизоцим, что препятствует преждевременному лизису бактериальных клеток после индукции профага и позволяет сконцентрировать бактериальные клетки перед получением упаковочных экстрактов. И, наконец, бактериальные лизогенные клетки содержат мутацию *gcsA*, которая предотвращает гомологичную рекомбинацию между ДНК профага и рекомбинантными ДНК, упаковываемыми в фаговые частицы.

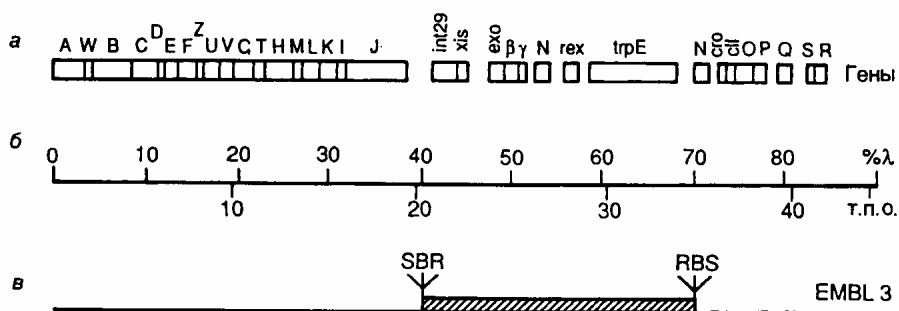


Рис. II.7. Генетическая карта хромосомы бактериофага λ -EMBL3

а – расположение генов на хромосоме; **б** – шкала длины хромосомной ДНК в процентах от длины λ -ДНК и т.п.о.; **в** – участок генома, замещаемый на клонируемый фрагмент ДНК соответствующего размера. S, B и R – сайты рестрикции *Sal*GI, *Bam*HI и *Eco*RI соответственно

В качестве примера рассмотрим генетическую карту векторной ДНК бактериофага λ -EMBL и кратко обсудим возможности этого вектора (рис. II.7). Векторы серии EMBL являются производными ДНК бактериофага λ 1059. Их хромосомная ДНК длиной в 42364 п.о. содержит центральный сегмент ДНК

длиной ~15 т.п.о., который замещается на клонируемый фрагмент ДНК соответствующего размера. При этом в фаговые частицы может быть упакована рекомбинантная ДНК общей длиной в 9–23 т.п.о. Замещаемый фрагмент фаговой хромосомы фланкирован с обоих концов последовательностями полилинкера, содержащего рестриктазные сайты *EcoRI*, *BamHI* и *SalGI*, по которым встраивают клонируемые фрагменты ДНК. При этом во время подготовки вектора к работе нет необходимости отделять "плечи" вектора от центрального фрагмента. Сначала центральный фрагмент ДНК выщепляется рестриктазой по одному из сайтов полилинкера, а затем смесь образовавшихся фрагментов обрабатывается другой рестриктазой, сайт для которой находится в полилинкере. Образующиеся олигонуклеотидные фрагменты полилинкера удаляются при переосаждении ДНК спиртом, а "липкие" концы "плеч" вектора и центральной последовательности получают некомплементарными друг другу и не могут объединяться в процессе лигирования с образованием исходной формы ДНК вектора.

7.2.3. Космиды и фазмиды

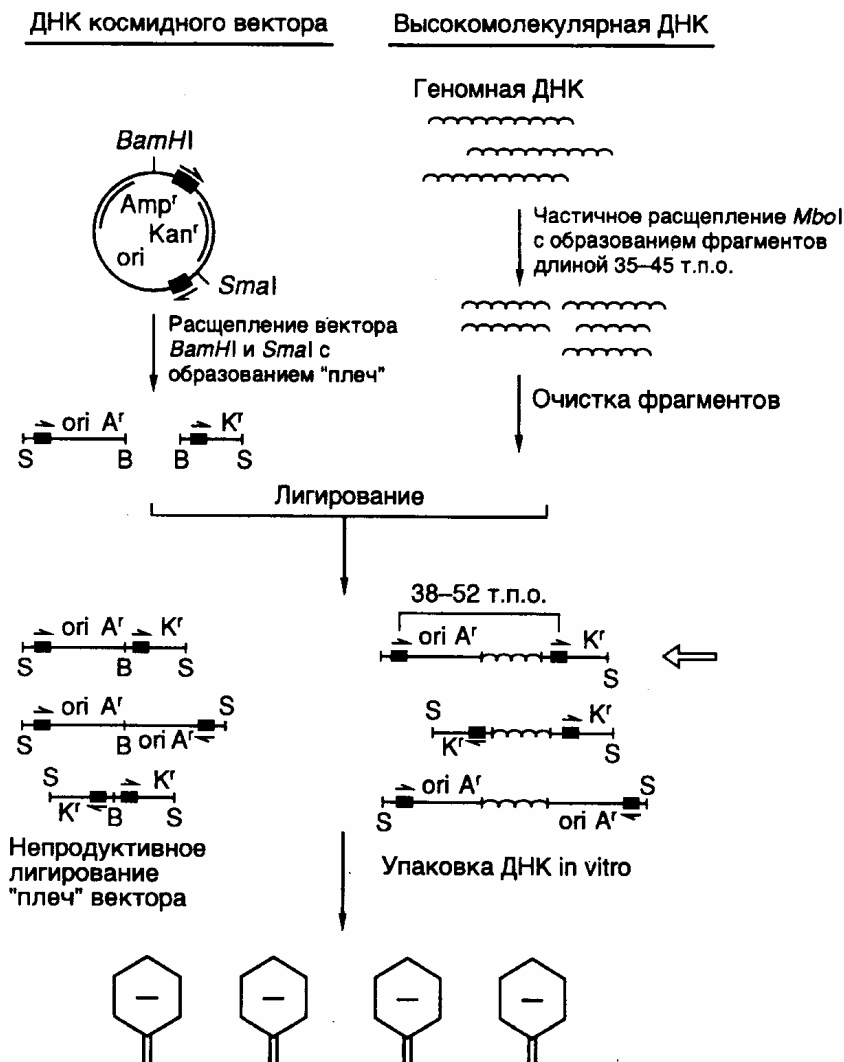


Рис. II.8. Космидный вектор и конструирование клонотеки геномной ДНК на его основе

Как уже упоминалось выше, фаговые векторы позволяют клонировать фрагменты ДНК длиной 15–25 т.п.о. Однако этого явно недостаточно, чтобы клонировать целиком многие гены животных и растений, длина которых зачастую превышает 35–40 т.п.о. Требуемой емкостью обладают векторные молекулы, называемые *космидами* (рис. II.8). Космиды представляют собой небольшие плазмиды, в которые in vitro введены *cos*-сайты ДНК фага λ . Отсюда

происходит название всего типа данных векторов (cosmid). В ДНК нормальных фаговых частиц *cos*-сайты расположены на концах молекул, они разделяют мономеры фаговой ДНК в *конкатемерах*, объединяющих несколько соединенных "голова к хвосту" мономеров, которые являются предшественниками зрелых фаговых ДНК перед упаковкой в фаговые частицы. В таких конкатемерах соседние *cos*-сайты располагаются на расстоянии 35–45 т.п.о. друг от друга и заключают между собой весь фаговый геном. В процессе упаковки *cos*-сайты узнаются компонентами ферментативной системы и по ним происходит последовательное отделение (отрезание) упакованной в фаговую частицу λ -ДНК от остальной неупакованной ДНК конкатемера. Таким образом, наличие *cos*-сайтов в ДНК является, по существу, единственным необходимым условием упаковываемости ДНК в фаговые частицы. Это означает, что последовательность нуклеотидов λ -ДНК, расположенная между двумя *cos*-сайтами, которая заключает в себе весь фаговый геном (35–45 т.п.о.), может быть замещена *in vitro* на аналогичный по длине (38–52 т.п.о.) фрагмент чужеродной ДНК и эффективно упакована в фаговые частицы (такова максимальная емкость головки фага). Естественно, что такая искусственная фаговая частица оказывается нежизнеспособной.

Однако после адсорбции химерной фаговой частицы на поверхности бактериальной клетки заключенная в ней ДНК проникает (вводится фаговой частицей) внутрь бактерии и начинает автономно реплицироваться как плазида, размер которой составляет 30–40 т.п.о. Поскольку такая плазида (космида) содержит в своем составе селективируемые маркеры в виде генов устойчивости к антибиотикам, ее поддерживают в бактериальных клетках путем выращивания бактерий на среде с соответствующими антибиотиками. Несмотря на то что емкость космидных векторов значительно выше фаговых, эффективность клонирования в космидах ниже, хотя и достигает в ряде случаев 10^5 – 10^6 колоний на 1 мкг клонируемой ДНК. При такой эффективности упаковки требуется всего лишь 2–4 мкг клонируемой ДНК для получения полной клонотеки большинства эукариотических геномов.

Стадия упаковки ДНК космид в фаговые частицы используется лишь для облегчения процесса введения рекомбинантных ДНК большого размера внутрь бактериальных клеток. Такой процесс имитирует проникновение фаговой хромосомы в бактерии во время фаговой инфекции. В случае космид сходство

между их проникновением в бактериальные клетки и фаговой инфекцией на этом заканчивается. Однако сходство является более глубоким в случае векторов, называемых *фазмидами*. Фазмиды представляют собой векторные молекулы ДНК, которые содержат в себе генетические элементы плазмид и хромосом бактериофагов. Они могут обладать емкостью в отношении клонируемой ДНК, характерной для λ -векторов, и существовать в определенных условиях в бактериальных клетках в виде плазмиды или же упаковываться в фаговые частицы *in vivo* при изменении этих условий.

7.2.4. Сверхъемкие векторы YAC, BAC и PAC

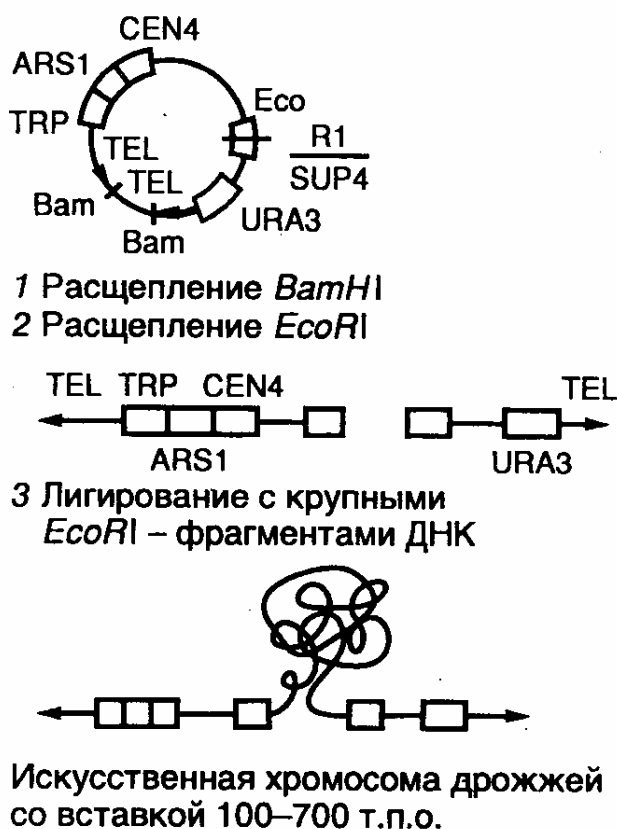


Рис. II.9. Схема клонирования сверхдлинных молекул ДНК с использованием вектора YAC

1 – линейаризация ДНК вектора рестриктазой *Bam*HI;
2 – расщепление линейаризованной ДНК вектора рестриктазой *Eco*RI с образованием "плечей"; 3 – введение в вектор клонируемого *Eco*RI-фрагмента ДНК

Хромосомы высших организмов содержат в своем составе протяженные молекулы ДНК. Например, длина ДНК одной из типичных хромосом человека составляет 100–200 миллионов пар оснований (м.п.о.). Исследование генов в хромосомах высших растений, животных и человека потребовало создания векторов для клонирования фрагментов ДНК длиной в несколько сотен тысяч пар оснований. Этим задачам отвечает недавно созданная система для клонирования сверхдлинных молекул ДНК на основе искусственно полученной мини-хромосомы дрожжей YAC (yeast artificial chromosome). YAC-вектор представляет собой кольцевую молекулу ДНК, содержащую ряд генетических элементов, которые позволяют ей существовать во внехромосомном состоянии в клетках дрожжей (рис. II.9).

Вектор включает в себе две теломерные последовательности нуклеотидов *TEL*, необходимые для репликации концов мини-хромосомы, и область начала репликации *ARS1*, соединенную с последовательностью центромеры. Все эти функциональные элементы требуются для репликации YAC-вектора и его правильной передачи в дочерние ядра во время митоза. Кроме того, вектор содержит два селектируемых маркера *TRP*, восстанавливающих способность к росту ауксотрофных по триптофану клеток дрожжей в отсутствие экзогенного триптофана, а также маркер *URA3*, компенсирующий генетический дефект клеток дрожжей, который нарушает биосинтез урацила. В векторе имеется также ген супрессорной тРНК *sup4*, являющийся селектируемым маркером для поддержания вектора в мутантных бактериальных клетках, содержащих амбер-мутации в жизненно важных генах. Помимо этого, он обладает последовательностями нуклеотидов, необходимыми для его репликации в бактериальных клетках.

При подготовке к клонированию YAC-вектор, выделенный в виде плазмиды, расщепляют рестриктазой *BamHI* и отделяют от образовавшегося короткого фрагмента ДНК, который не требуется для репликации YAC-вектора в дрожжах (этап 1). После этого проводят второе расщепление вектора рестриктазой *EcoRI*, сопровождающееся образованием двух его "плеч", каждое из которых на одном из концов содержит теломерные последовательности хромосомы дрожжей (этап 2). На заключительном этапе (3) полученные "плечи" лигируют с крупными *EcoRI*-фрагментами клонируемой ДНК, которые получают путем частичного расщепления высокомолекулярной хромосомной ДНК,

предназначенной для клонирования. Полученными таким образом рекомбинантными ДНК трансформируют протопласты клеток дрожжей, и образовавшиеся трансформанты отбирают на селективной твердой питательной среде. В таком векторе удавалось осуществлять клонирование фрагментов ДНК длиной до 700 т.п.о.

При всех своих достоинствах системы клонирования, основанные на векторах семейства YAC, обладают рядом существенных недостатков. В рекомбинантных ДНК, поддерживаемых в таких системах, часто возникают внутренние делеции. Кроме того, при введении рекомбинантных ДНК в клетки дрожжей иногда имеет место проникновение в одну клетку нескольких молекул вектора со вставками. В итоге отдельные клоны дрожжевых клеток могут содержать несколько несцепленных друг с другом молекул рекомбинантных ДНК, а рекомбинация между ними вообще может приводить к образованию химерных молекул. Все это очень затрудняет физическое картирование генов в хромосомах исследуемых объектов. Для преодоления такого рода трудностей были сконструированы альтернативные векторные системы, среди которых наиболее популярными в настоящее время являются системы, основанные на *искусственных хромосомах бактерий* – BAC (bacterial artificial chromosome).

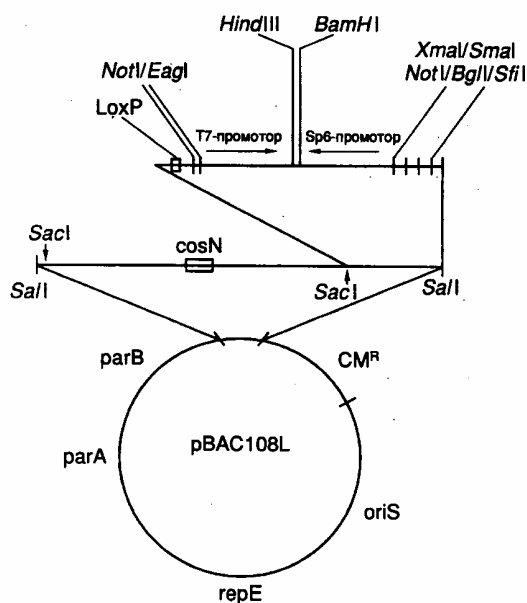


Рис. II.10. Вектор pBAC108L – представитель семейства BAC-векторов

В векторных системах ВАС используется ДНК хорошо изученного *полового фактора (F-фактора)* *E. coli* – гигантской плазмиды мужских бактериальных клеток, которые являются донорами бактериальной ДНК при конъюгации с женскими клетками (рис. II.10). Типичный F-фактор содержит гены *oriS*, *repE*, *parA* и *parB*, регулирующие его собственную репликацию и контролирующие число его копий в бактериальных клетках. В частности, гены *oriS* и *repE* обеспечивают однонаправленную репликацию F-фактора, а гены *parA* и *parB* поддерживают число его копий на уровне одной-двух на бактериальную клетку. Классический вектор ВАС (pBAC108L) включает в себя все эти гены, а также ген устойчивости к хлорамфениколу, используемый в качестве селектируемого маркера. Вектор содержит также сегмент ДНК, по которому производится клонирование. В этом сегменте имеется типичный полилинкер, а также два уникальных сайта рестрикции *HindIII* и *BamHI*, фланкированные промоторами T7- и Sp6-ПНК-полимераз. Эти промоторы могут быть использованы для получения РНК-зондов, необходимых для осуществления "прогулок по хромосомам" (раздел 7.6.2), а также прямого секвенирования клонированной ДНК в месте стыковки с вектором. Кроме того, во фрагменте имеется сайт *cosN*, обеспечивающий расщепление вектора со вставкой в уникальном месте с помощью *терминазы* фага λ без применения ферментов рестрикции. Этот фермент используется бактериофагом для специфического разрезания конкатемеров своей хромосомы при упаковке в фаговые частицы. Для той же цели может быть применен и сайт *loxP* бактериофага P1, который является мишенью для фаговой эндонуклеазы Cre. Эти сайты используются для получения специфических концов клонированной ДНК с целью ее дальнейшего рестрикционного картирования путем введения концевой метки с последующим неполным расщеплением с помощью рестриктаз и электрофоретическим разделением образовавшихся фрагментов (см. раздел 7.6.1).

Современные ВАС-векторы позволяют клонировать фрагменты ДНК длиной до 300 т.п.о. и выше. Рекомбинантные молекулы вводятся в клетки *E. coli* с помощью электропорации (см. раздел 7.3.2), причем эффективность образования трансформантов в 10–100 раз выше, чем при обычной трансформации сферопластов дрожжей векторами семейства YAC. Это позволяет уменьшить исходное количество ДНК, необходимое для

конструирования репрезентативных клонотек генов (см. раздел 7.3). При скрининге таких клонотек используются традиционные методы работы с бактериальными колониями (раздел 7.3.3). В отличие от YAC-ДНК, которая находится в клетках дрожжей в линейной форме, BAC-векторы со вставками, как и традиционные F'-факторы, существуют в бактериальных клетках в виде кольцевых суперскрученных молекул. Это облегчает их выделение и последующую работу с рекомбинантными молекулами ДНК в растворе, а кроме того, допускает простое повторное введение в бактериальные клетки этих ДНК, выделенных минипрепаративными методами. Поскольку рекомбинантные BAC-векторы существуют в бактериальных клетках в виде одной копии, исключаются совместное клонирование в одной клетке разных фрагментов ДНК и образование химерных молекул, что весьма существенно для физического картирования больших геномов методами "снизу вверх" (см. раздел 12.1.4). Весьма существенным свойством системы клонирования, основанной на векторах семейства BAC, является ее генетическая стабильность. Исходная структура клонированных фрагментов ДНК в пределах точности использованных методов сохраняется в таких векторах даже после 100 серийных пересевов бактериальных клеток, содержащих рекомбинантные молекулы ДНК. Все вышеперечисленные свойства переводят векторы BAC в разряд сверхъёмких векторов нового поколения.

В заключение следует упомянуть о семействе векторов PAC (P1-derived artificial chromosome), также часто используемых в современных исследованиях. Векторы этой серии содержат гены бактериофага P1, обеспечивающие репликацию фаговой хромосомы в зараженных бактериальных клетках. Рекомбинантные ДНК на их основе (размер вставки 150–200 т.п.о.) также вводятся в бактериальные клетки с помощью электропорации.

7.2.5. Интегрирующие и челночные (бинарные) векторы

Векторы, пригодные для клонирования ДНК в бактериях, отличающихся от *E. coli*, должны обладать всеми характерными чертами, которые были отмечены выше. От только что рассмотренных они отличаются главным образом тем, что содержат природные или искусственные генетические элементы, функционирующие в новых клетках-хозяевах.

Интегрирующие векторы грамположительной бактерии *Bacillus subtilis*. *B. subtilis*, как и *E. coli*, является излюбленным объектом генной инженерии. Это связано с тем, что *B. subtilis* – непатогенный микроорганизм, многие штаммы которого широко используются в микробиологической промышленности для производства биологически активных соединений и пищевых веществ. В отличие от *E. coli* *B. subtilis* способна секретировать белки и пептиды, что облегчает их очистку и дальнейшее использование. Большинство векторов для клонирования ДНК в клетках *B. subtilis* создано на основе плазмид других видов бацилл, а также *Staphylococcus aureus* и бактерий рода *Streptococcus*.

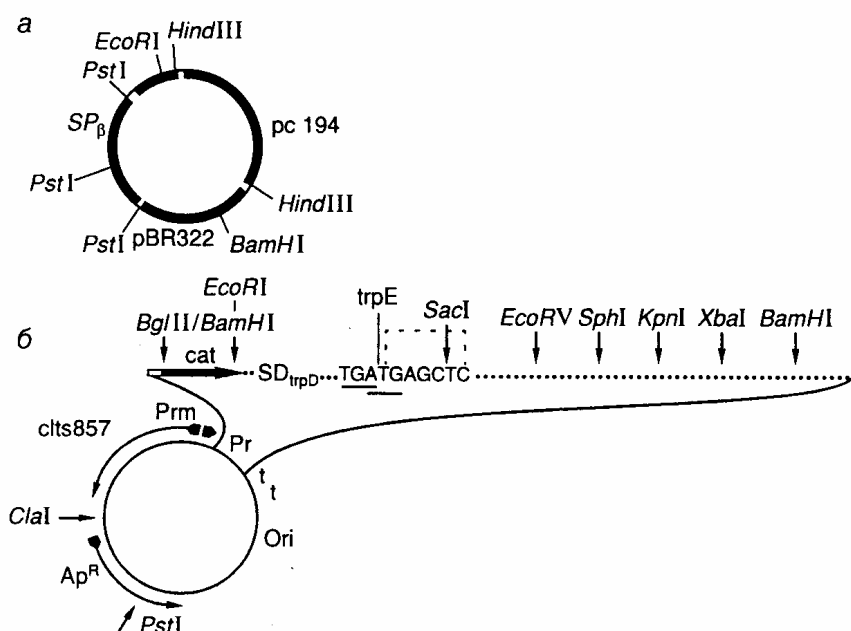


Рис. II.11. Челночный вектор pFH7 (а) и экспрессирующий вектор pPR-TGATG-1 (б)

Обозначено расположение генов и уникальных сайтов рестрикции

Основным свойством интегрирующих векторов является их способность стабильно встраиваться в геном клетки-хозяина. Это становится возможным благодаря наличию в таких векторах последовательностей нуклеотидов, гомологичных последовательностям геномной ДНК. В результате функционирования общей системы рекомбинации происходит объединение хромосомной и плазмидной ДНК интегрирующего вектора, которое приводит к стабильному включению всей векторной плазмиды в хромосому. Примером

такого интегрирующего вектора служит плаزمида pFH7 *B. subtilis* (рис. II.11,а). Векторная плазмида содержит фрагмент ДНК умеренного бактериофага SPβ и после попадания в клетки *B. subtilis*, лизогенные по данному бактериофагу, эффективно интегрируется в профаг. Поскольку плазмида содержит ген устойчивости к хлорамфениколу *cat*, клетки приобретают этот признак. Индукция профага приводит к образованию фаговых частиц, трансдуцирующих такую плазмиду и ассоциированный с ней признак устойчивости к хлорамфениколу. Интеграция плазмиды SPβ в бактериальную хромосому происходит по механизму гомологичной рекомбинации с участием гена *recE*.

Способность к интеграции в бактериальную хромосому обнаруживают и другие плазмиды, содержащие фрагменты хромосомной ДНК клеток-хозяев, что продемонстрировано, в частности, для плазмид *E. coli* и *Streptococcus pneumoniae*.

Интегрирующие векторные системы, в которых используется тот же принцип гомологичной рекомбинации, разработаны и для эукариотических клеток, включая клетки животных и растений. В конце концов, такие работы привели к развитию целого направления исследований по созданию трансгенных животных и растений, стабильно наследующих и экспрессирующих гены, искусственно введенные в их геном. О некоторых важных следствиях этого направления исследований, включая генотерапию, речь пойдет в главе 10.

Наличие феномена гомологичной рекомбинации между хромосомной ДНК клетки-хозяина и векторной ДНК, содержащей гомологичные хромосомной ДНК последовательности нуклеотидов, приходится учитывать при получении соответствующих генно-инженерных конструкций. Такая неконтролируемая рекомбинация в большинстве случаев нежелательна, так как может приводить к потере или структурным перестройкам клонированных фрагментов ДНК. Для того чтобы свести последствия этого явления к минимуму, используют специальные штаммы клеток-хозяев, в которых общая система рекомбинации не функционирует, например вследствие мутационной инактивации гена *recA* *E. coli* или *recE* *B. subtilis*.

Челночные векторы. Интегрирующая плазмида pFH7 (см. рис. II.11,а) дает возможность проиллюстрировать еще один важный принцип, широко используемый при конструировании векторных систем в генной инженерии. Эта

плазмида получена путем объединения двух репликонов, один из которых берет начало от плазмиды pC194 *B. subtilis*, а другой – от плазмиды pBR322 *E. coli*, что дает возможность вектору существовать и стабильно реплицироваться как в клетках *E. coli*, так и *B. subtilis*. Такие векторы, способные реплицироваться в клетках-хозяевах разных биологических видов, называют *челночными*, или *бинарными* векторами.

Принципы конструирования и функционирования челночных векторов одинаковы, они должны включать в себя репликоны тех генетических систем, в которых будет происходить репликация челночного вектора. При этом используются области начала репликации генетических элементов, которые автономно существуют во внехромосомном состоянии в природных условиях. Так, обсуждавшийся выше интегрирующий вектор pFH7 *B. subtilis* обладает свойствами челночного вектора, поскольку для его конструирования использованы репликоны двух видов бактерий. Более эффективными примерами челночных векторов являются плазмидные ДНК, способные реплицироваться в клетках высших (животных и растений) и низших организмов. Необходимость использования челночных векторов в геной инженерии связана с тем, что наработку в препаративном количестве векторной ДНК для проведения гено-инженерных манипуляций удобнее проводить в бактериальных клетках, тогда как получение биологически активных продуктов клонированных генов высших организмов во многих случаях возможно только в клетках своего или близкого вида, в которых эти гены экспрессируются в природных условиях, т.е. в своем обычном генетическом окружении (подробнее см. раздел 7.6).

7.2.6. Конструирование экспрессирующих векторов и их функционирование

Первая часть книги была посвящена описанию механизмов, обеспечивающих высокоэффективную и высокоспецифическую экспрессию генов. Такого рода информацию успешно используют в настоящее время для эффективной экспрессии рекомбинантных генов в гомологичном или гетерологичном генетическом окружении. После рассмотрения основных принципов конструирования векторов для клонирования ДНК можно перейти к обсуждению проблемы экспрессии клонированных генов в искусственных генетических системах. Именно экспрессия клонированных генов является одной из основных задач геной инженерии. Действительно, функциональную

роль отдельных генов и их частей в живом организме можно понять и оценить лишь по их экспрессии, т.е. по фенотипическому проявлению их потенциальных биологических возможностей. Для получения полноценной экспрессии клонированных генов используют экспрессирующие векторные системы, принципы конструирования которых в настоящее время хорошо разработаны.

Напомним, что экспрессия генетической информации, заключенной в генах живых организмов, в соответствии с основным постулатом молекулярной биологии осуществляется путем ее переноса от нуклеиновых кислот к белкам по схеме: ДНК $\leftrightarrow$ РНК $\rightarrow$ белок. Любой полноценный в функциональном отношении ген, в том числе и рекомбинантный, состоит из двух основных частей – регуляторной и структурной. Генетический код, с помощью которого последовательность нуклеотидов в кодирующей структурной части гена определяет последовательность аминокислот в кодируемом этим геном белке, за немногими исключениями, универсален для всех живых организмов. Это дает принципиальную возможность получения полноценной экспрессии самых разнообразных осмысленных последовательностей нуклеотидов, в том числе и искусственно синтезированных, в любом природном генетическом окружении. Таким образом, универсальность генетического кода является прочной основой для создания функционально активных генно-инженерных конструкций.

Регуляторная часть генов распознается соответствующими ферментными системами организма и обеспечивает упорядоченную экспрессию его структурной части. При этом регуляторная часть гена, как правило, высокоспецифична в отношении своих природных генетических эффекторов (РНК-полимераз, рибосом, факторов транскрипции и трансляции, белковых факторов сплайсинга, ферментов, осуществляющих посттрансляционные модификации полипептидов, и т.п.), и чаще всего она не может эффективно функционировать в гетерологичном генетическом окружении. Очевидно, что при конструировании высокоэффективных экспрессирующих векторов необходимо, прежде всего, учитывать особенности структуры регуляторной части рекомбинантного гена, исходя из того, в каких генетических условиях клонированный ген предполагается экспрессировать. Однако не только регуляторные последовательности генов являются препятствием для высокоэффективной экспрессии чужеродных рекомбинантных генов. Структурные части генов про- и эукариот

фундаментально отличаются друг от друга по наличию у последних внутри генов многочисленных и протяженных некодирующих последовательностей нуклеотидов – интронов. Следовательно, гены эукариот не могут эффективно экспрессироваться в клетках прокариот, поскольку у прокариот отсутствуют соответствующие системы сплайсинга. Кроме того, у предшественников эукариотических мРНК не может осуществиться в бактериальных клетках и правильный процессинг 3'- и 5'-концевых некодирующих последовательностей. Даже такой беглый взгляд на проблему экспрессии рекомбинантных генов в чужеродном генетическом окружении позволяет увидеть многочисленные препятствия, лежащие на пути решения этой проблемы. Ниже будут рассмотрены факторы, которые необходимо учитывать при конструировании экспрессирующих векторов.

Факторы, оказывающие влияние на эффективность экспрессии рекомбинантных генов: транскрипция. Для того чтобы в клетках *E. coli* или любого другого организма произошла экспрессия клонированного гена, этот ген должен быть транскрибирован РНК-полимеразой, а образовавшаяся РНК транслирована рибосомами с образованием нативного белкового продукта. Следовательно, для эффективной экспрессии в клетках *E. coli* любой ген должен находиться под контролем регуляторных элементов данного микроорганизма. Особенности структуры и функционирования регуляторных элементов генов про- и эукариот были подробно рассмотрены в первой части книги. Знание основных принципов строения и функционирования промоторов позволяет конструировать новые промоторы, которые по своей эффективности могут превосходить природные. Примером такого промотора, используемого в некоторых экспрессирующих векторах, является промотор *Taq*, который в своем составе объединяет последовательности промоторов лактозного и триптофанового оперонов *E. coli*.

Базальный уровень транскрипции рекомбинантных генов гарантируется небольшим набором упорядоченных регуляторных последовательностей. Например, у *E. coli* это происходит при наличии перед структурной частью клонированного гена последовательностей в окрестностях нуклеотидов +1, -10 и -35.

a.....tcTTGACat..t.....t.tg.TAtAat.....cat

Однако даже у *E. coli* транскрипция большинства генов носит регулируемый характер, что обеспечивается системой белков-репрессоров и активаторов, а также низкомолекулярных эффекторов, которые главным образом влияют на взаимодействие РНК-полимеразы с транскрибируемыми генами. Необходимость регулирования экспрессии генов в генной инженерии диктуется, прежде всего, тем, что сверхэкспрессия клонированных генов в клетках-продуцентах рекомбинантных белков при использовании сильных конститутивных промоторов может отрицательно сказываться на их жизнеспособности, поскольку промежуточные и конечные продукты экспрессии часто обладают цитотоксическим действием. В этом случае сверхэкспрессия клонированных генов сопровождается снижением скорости или даже полной остановкой роста бактериальных или иных клеток. Для предотвращения такого рода эффектов в генной инженерии и биотехнологии используют генно-инженерные конструкции, в которых экспрессия клонированных генов в значительной степени подавлена (репрессирована) на ранних фазах роста культуры рекомбинантных клеток и может быть дерепрессирована в любой нужный момент времени.

В качестве примера рассмотрим две системы регулируемой экспрессии рекомбинантных генов в клетках *E. coli*, которые получили широкое распространение. В одном случае для контроля экспрессии клонированных генов используют хорошо изученную систему регуляторных элементов лактозного оперона, в другом – систему промотор–репрессор–оператор фага λ . Принцип регуляции экспрессии рекомбинантных генов в обоих случаях один и тот же. В векторные молекулы вводятся регуляторные последовательности фага λ или *lac*-оперона, за которыми следуют последовательности нуклеотидов полилинкера с несколькими уникальными сайтами рестрикции, по которым проводится клонирование исследуемого гена. Одновременно в тот же вектор вводится ген-регулятор, кодирующий белок-репрессор. В отсутствие индуцирующего воздействия молекулы репрессора, ген которого экспрессируется конститутивно, связываются с оператором, тем самым препятствуя взаимодействию РНК-полимеразы с промотором, под контролем которого находится клонированный рекомбинантный ген. Это приводит к

резкому снижению уровня транскрипции клонированного гена, низкому внутриклеточному содержанию соответствующей мРНК и белкового продукта ее трансляции.

Существенным различием двух систем регулируемой экспрессии является способ их индукции. В случае регуляторной системы *lac*-оперона чаще всего используют синтетический аналог природного индуктора (*алло*-лактозы) – изопропил- β -D-тиогалактопиранозид (IPTG). Индуктор связывается с молекулой репрессора и нарушает его взаимодействие с операторным участком ДНК. Область *lac*-промотора становится доступной для РНК-полимеразы, которая с высокой эффективностью начинает транскрибировать гены, расположенные вслед за промотором. В отличие от этого в системе репрессии-дерепрессии фага λ используют ген репрессора *ci*, содержащий температурочувствительную мутацию. Наличие такой мутации приводит к тому, что мутантный белок-репрессор сохраняет свою нативную конформацию, а следовательно, и способность подавлять синтез РНК только при перmissive температуре (как правило, 28–30°). При повышении температуры окружающей среды до 42° происходят инактивация белка-репрессора и дерепрессия генов, расположенных под контролем промоторов P_L или P_R фага λ .

Использование системы репрессии-дерепрессии, чувствительной к температуре, более технологично по сравнению с химической индукцией по двум обстоятельствам: во-первых, это менее дорогой способ индукции; во-вторых, для системы фага λ характерна более высокая степень репрессии генов, находящихся под ее контролем, и соответственно более низкий базальный уровень их экспрессии. Такой фактор становится решающим при клонировании и получении экспрессии генов, белковые продукты которых токсичны для клетки-хозяина. Однако несмотря на технологичность системы репрессии-дерепрессии, чувствительной к температуре, сверхпродукция рекомбинантных белков, особенно эукариотических, часто сопровождается внутриклеточным образованием их нерастворимых агрегатов, что в ряде случаев полностью отсутствует при использовании для выращивания рекомбинантных клеток температур ниже 30° (подробнее см. ниже). Эти эффекты накладывают ограничения на использование векторов с температурочувствительным контролем экспрессии для продукции

эукариотических и в меньшей степени прокариотических белков в бактериальных клетках.

Для обеспечения регулируемой тканеспецифической экспрессии рекомбинантных генов в соматических клетках животных и растений в составе векторов используют энхансеры, которые избирательно стимулируют транскрипцию в соответствующих тканях и не оказывают такого действия на гены в тканях, клетки которых не экспрессируют необходимые регуляторные белки. Кроме того, популярным становится введение в экспрессирующие эукариотические векторы пограничных последовательностей нуклеотидов, фланкирующих клонируемые гены, которые помогают обеспечивать экспрессию рекомбинантных генов, сводя к минимуму эффект их положения в хромосомах соматических клеток. О механизмах этих явлений подробно говорилось в первой части книги.

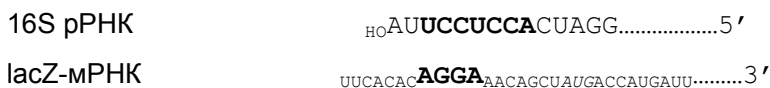
В настоящее время, по-видимому, не существует систем репрессии клонируемых генов, которые бы полностью блокировали их транскрипцию. Это создает определенные проблемы при клонировании генов белков, токсичных для клеток-хозяев, таких как колицины, эндо- и экзонуклеазы, протеиназы и т.п.

Наряду с промоторами векторы, экспрессирующие рекомбинантные гены в клетках *E. coli*, должны содержать и гомологичные терминаторы транскрипции. Необходимость в этом вызвана тем, что даже нетранслируемые избыточные 3'-концевые последовательности нуклеотидов мРНК, транскрибированной с рекомбинантного гена, отрицательно сказываются на эффективности ее трансляции, что, по-видимому, связано с участием таких последовательностей в образовании сложной третичной структуры рекомбинантных мРНК.

Факторы, влияющие на эффективность экспрессии рекомбинантных генов: трансляция. Как следует из изложенного выше, эффективность экспрессии клонированных генов в бактериальных и эукариотических клетках во многом зависит от эффективности их транскрипции, т.е. от уровня внутриклеточного содержания соответствующих мРНК. Однако эффективность экспрессии рекомбинантных генов не в меньшей степени зависит и от эффективности трансляции таких мРНК рибосомами. При работе с рекомбинантными ДНК чаще всего приходится иметь дело с гетерологичными системами, в которых клонированные экспрессируемые ДНК чужеродны по

отношению к генетическому окружению и ферментативным системам бактериальных или эукариотических клеток-хозяев. В первой части книги в качестве примеров были рассмотрены некоторые факторы, играющие большую роль в регуляции экспрессии генов на посттранскрипционном уровне. Такие факторы вносят свой вклад и в эффективность экспрессии рекомбинантных генов. Однако в этом случае необходимо всегда иметь в виду уровень соответствия регуляторных факторов клеток-хозяев регуляторным последовательностям мРНК, которые являются транскриптами рекомбинантных генов. В частности, на эффективность трансляции чужеродных мРНК в бактериальных клетках оказывает сильное влияние первичная структура небольшого участка 5'-концевой части мРНК перед иницирующим AUG-кодоном, частично комплементарная 3'-концевой части 16S рРНК. Этот участок мРНК называют последовательностью Шайна–Дальгарно, или SD-последовательностью.

SD-последовательность мРНК служит сигналом для взаимодействия рибосом с мРНК и обеспечивает инициацию трансляции с правильного AUG-кодона. Она представляет собой короткую последовательность нуклеотидов, обогащенную пуриновыми основаниями, и расположена за 4–9 нуклеотидов перед AUG-кодоном:



(в приведенном примере прописными буквами у lacZ-мРНК изображена SD-последовательность нуклеотидов, комплементарная соответствующей пиримидиновой последовательности 16S рРНК, и подчеркнут иницирующий кодон). Различия в эффективности распознавания такого сайта на мРНК рибосомами приводят к различиям в эффективности экспрессии соответствующих генов. Достаточно наличия в SD-сайте трех нуклеотидов, комплементарных соответствующему участку 16S рРНК, чтобы рибосомы приобретали способность связывать эту мРНК. Помимо SD-последовательности в мРНК, по-видимому, имеются и другие участки, необходимые для эффективной инициации трансляции рибосомами, так как на основании только первичной структуры мРНК в настоящее время невозможно

предсказать эффективность сайта связывания рибосом. Более того, сайты инициации трансляции на мРНК, обладающие протяженной гомологией с 16S рРНК, являются малоэффективными, возможно, из-за слишком прочного связывания рибосом с мРНК.

Существенным фактором, оказывающим влияние на эффективность трансляции мРНК рибосомами, может быть вторичная структура мРНК в окрестностях иницирующего кодона. Из всего вышеизложенного следует, что некодирующая 5'-концевая часть мРНК играет важную роль в эффективности ее трансляции рибосомами, и, по крайней мере, эти особенности их трансляции необходимо учитывать при конструировании векторов для экспрессии рекомбинантных генов в клетках *E. coli*. Современные экспрессирующие векторы построены таким образом, что промотор, с которого начинается транскрипция клонированного гена, а также следующая за ним последовательность нуклеотидов, включающая иницирующий ATG-кодон, являются частью генов, эффективно экспрессирующихся в клетках *E. coli*. Сразу вслед за ATG-кодом вектора располагают уникальные сайты рестрикции, по которым проводят встраивание кодирующей части клонируемого гена, экспрессию которого хотят получить. В том случае, если сайт рестрикции, по которому производится клонирование, расположен в начале кодирующей части природного гена, в результате транскрипции такой рекомбинантной молекулы образуется химерная мРНК. 5'-Концевая часть этой мРНК от промотора до сайта рестрикции для клонирования соответствует последовательности нуклеотидов вектора и кодирует несколько аминокислот природного гена, регуляторные последовательности которого использовали для конструирования экспрессирующего вектора. Протяженная же 3'-концевая часть такой мРНК будет соответствовать клонированной последовательности нуклеотидов. Если соединение этих двух частей мРНК произошло в одной рамке считывания, то в результате ее трансляции будет образовываться химерный полипептид, N-концевая часть которого кодируется вектором, а С-концевая – клонированным фрагментом ДНК. Примерами таких экспрессирующих векторов могут быть плазмиды pUC18 и pUC19 (см. рис. II.5,б), продукты экспрессии которых представляют собой химерные белки, а N-концевая часть содержит N-концевые аминокислотные остатки β -галактозидазы *E. coli*.

Превосходным вектором, прекрасно зарекомендовавшим себя в работе, является экспрессирующий вектор *pPR-TGATG-1*, сконструированный в лаборатории С.В. Машко (г. Москва). Вектор функционирует по тому же самому принципу. В нем структурная часть экспрессируемого гена соединяется через уникальный сайт рестрикции *SacI* с иницирующим кодоном ATG (см. рис. II.11,б). В результате образуется гибридный оперон, транскрибируемый с промотора λ -*P_R*, дистальный цистрон которого представлен клонируемым геном, а цистрон, проксимальный промотору *P_R*, содержит гибридный ген *cro'-cat'-trpE*, состоящий из слитых частей генов *cro* фага λ , хлорамфениколацетилтрансферазы *cat* и *trpE* *E. coli*. 3'-Концевая часть цистрона содержит область реинициации трансляции, SD-последовательность гена *trpD* *E. coli* и перекрывающиеся кодоны терминации и инициации трансляции (TGATG). В этой последовательности 3'-концевой нуклеотид G является первым нуклеотидом уникального сайта рестрикции *SacI* (GAGCTC). В процессе синтеза белка, направляемого вектором, образуется короткий гибридный пептид Cro'-Cat'-TrpE длиной ~120 аминокислот. Сразу же после терминации его синтеза с высокой эффективностью происходит реинициация синтеза полипептида, кодируемого клонированным геном, на перекрывающемся с кодоном TGA ATG-кодоне. Поскольку промотор *P_R* в этом векторе контролируется температурочувствительным репрессором фага λ *clt857*, экспрессия клонированного гена индуцируется повышением температуры окружающей среды до 42° и сопровождается инактивацией репрессора.

Приведенные примеры четко демонстрируют роль двух основных факторов, оказывающих влияние на эффективность экспрессии чужеродных клонированных генов, находящихся в составе экспрессирующих векторов в бактериальных клетках. Этими факторами являются, во-первых, оптимальная транскрипция клонированных генов и, во-вторых, эффективная трансляция транскрибированной мРНК. В обоих случаях наилучшие результаты могут быть получены при объединении в составе вектора структурной части клонированного чужеродного гена с генетическими элементами, регулирующими транскрипцию и трансляцию (промоторы, SD-последовательности и т.п.) бактериальных клеток-хозяев или их вирусов.

Знание главных принципов, лежащих в основе регуляции транскрипции и трансляции, позволяет конструировать новые регуляторные последовательности (например промотор *Taq*) путем объединения известных регуляторных элементов или использования искусственно синтезированных последовательностей нуклеотидов, представляющих собой усредненные канонические регуляторные последовательности, учитывающие особенности большого числа природных регуляторных элементов. Однако этими важными факторами не исчерпывается возможность оптимизации экспрессии рекомбинантных генов в чужеродном генетическом окружении.

Среди других факторов, оказывающих влияние на эффективность трансляции, следует упомянуть частоту встречаемости кодонов при кодировании белков в структурных частях разных генов. В настоящее время установлено, что использование синонимических кодонов (кодирующих одну и ту же аминокислоту) вырожденного генетического кода не случайно и отражает количественную представленность отдельных изоакцепторных тРНК в клетках организма. С другой стороны, частота использования кодонов в разных генах одного и того же организма является эффективным фактором, регулирующим уровень экспрессии этих генов в процессе трансляции. Чем реже тот или иной кодон используется для кодирования аминокислоты в данном организме, тем с меньшей скоростью он будет транслироваться рибосомами вследствие низкой внутриклеточной концентрации тРНК, узнающей такой кодон. Так, в группе белков дрожжей с высоким уровнем экспрессии 96% аминокислот кодируются только 25 из 61 возможного кодона. Замена этих 25 предпочтительных кодонов в 5'-концевых частях генов белков с высоким уровнем экспрессии на редко встречающиеся кодоны снижает уровень синтеза как белков, так и их мРНК.

Кроме того, на эффективность трансляции кодона оказывает влияние и контекст кодона, т.е. соседние, окружающие кодон последовательности нуклеотидов. Этот феномен предпочтительного использования одних кодонов перед другими необходимо учитывать в опытах по оптимизации экспрессии чужеродного генетического материала в гетерологичном генетическом окружении, поскольку несоответствие в частотах использования кодонов у организмов разных видов может служить серьезным препятствием для получения эффективной экспрессии клонируемых генов. Снижение уровня экспрессии под влиянием подобных факторов, в частности, наблюдали при

попытках экспрессии генов растений в клетках дрожжей, а также в других гетерологичных системах. Таким образом, эффекты, возникающие в связи с различной частотой использования кодонов, необходимо учитывать при подборе генетических систем для эффективной экспрессии клонированных чужеродных генов. Этот фактор может оказаться одним из решающих при получении полностью синтетических генов, например генов пептидных гормонов, для использования их в биотехнологии, а также при планировании мутационных замен нуклеотидов в белковой инженерии.

Факторы, влияющие на эффективность экспрессии рекомбинантных генов: стабильность рекомбинантных белков и РНК.

Важным фактором, определяющим уровень экспрессии рекомбинантных генов в реципиентных клетках, является стабильность кодируемых этими генами мРНК и белков. В первой части книги подробно рассматривалось значение таких факторов в регуляции экспрессии генов. Не менее важную роль эти феномены играют и в контроле уровня экспрессии рекомбинантных генов в искусственных генетических системах.

Как было упомянуто выше, замены кодонов в мРНК дрожжей на синонимические приводят не только к снижению уровня трансляции измененных мРНК в клетках дрожжей, но и сопровождаются уменьшением стабильности самих мРНК. Требование достаточного уровня стабильности мРНК является одним из факторов, определяющих давление отбора на использование конкретных синонимических кодонов для кодирования белков. Важную роль во внутриклеточной стабильности мРНК играют элементы ее вторичной структуры. Так, образование шпилек на 3'-конце мРНК препятствует ее деградации под действием бактериальных экзонуклеаз, РНКазы II и полинуклеотидфосфорилазы. Поскольку изменение первичной структуры с учетом возможной вторичной структуры без изменения функциональных свойств представляется в настоящее время сложной задачей, проблема внутриклеточной стабильности рекомбинантных РНК решается главным образом путем введения мутаций в гены экзо- и эндонуклеаз бактериальной клетки-хозяина. Так называемые безРНКазные штаммы *E. coli* находят применение в качестве реципиентов рекомбинантных плазмид в генной инженерии. При этом решается двойная задача: с одной стороны, в таких клетках повышается стабильность транскриптов рекомбинантных генов, а с

другой – пониженный внутриклеточный уровень определенных РНКаз способствует более эффективной очистке рекомбинантных белков, так как для многих препаратов ферментов даже незначительные примеси нуклеазной активности бывают нежелательны.

Проблема внутриклеточной стабильности рекомбинантных белков больше связана с деградацией небольших пептидов, поскольку крупные нативные белки более стабильны в бактериальных клетках. Один из подходов, позволяющих стабилизировать короткие чужеродные пептиды в клетках *E. coli*, – включение требуемого пептида в состав гибридного белка. В этом случае последовательность нуклеотидов, кодирующую гибридный белок, соединяют в составе экспрессирующего вектора в одной рамке считывания с бактериальным геномом, кодирующим белок (например геном β -галактозидазы). Образующийся в результате экспрессии такого рекомбинантного гена гибридный белок в своем составе содержит в N- или C-концевой части требуемый пептид, защищенный основным белком-носителем от протеолитической деградации. Для отделения пептида от белка-носителя их соединяют друг с другом последовательностью аминокислот, по которой можно провести специфическое расщепление полипептидной цепи. В том случае, если пептиды не содержат метионина, соединение белка-носителя и пептида осуществляют через эту аминокислоту, и отщепление пептида производят с помощью бромистого циана. Такой подход был использован, например для получения рекомбинантного соматостатина, а также A- и B-цепей инсулина.

Для разделения рекомбинантного пептида и белка-носителя можно использовать и протеолитические ферменты. Для этого требуемый пептид соединяют с белком-носителем с помощью последовательности аминокислот, специфически расщепляемой протеолитическим ферментом, например пептидазой клостридий (кlostрипаином) в случае получения кальцитонина или трипсином для получения β -эндорфина. Второй подход, позволяющий защищать рекомбинантные белки от внутриклеточного протеолиза, – выведение их в периплазматическое пространство или окружающую среду после завершения биосинтеза. При таком подходе рекомбинантные полипептиды соединяют с сигнальными последовательностями аминокислот белков (β -лактамаза, OmpA, OmpF, PhoA), секретируемых во внешнюю среду и

периплазматическое пространство. Подобным путем были получены, например проинсулин, гормон роста человека и β -эндорфин.

Особенности биосинтеза рекомбинантных эукариотических белков в бактериальных клетках: тельца включения. Экспрессия эукариотических генов в бактериальных клетках связана еще с одной проблемой: большая часть рекомбинантных эукариотических белков, а также некоторые прокариотические белки при очень высоком уровне биосинтеза переходят в нерастворимое состояние, образуя в клетках так называемые *тельца включения*. Тельца включения представляют собой частицы, состоящие главным образом из агрегатов рекомбинантного белка и ряда бактериальных белков. Рекомбинантные белки в тельцах включения находятся, как правило, в денатурированном неактивном состоянии, и для их получения в растворимом виде приходится применять мощные денатурирующие агенты, такие как мочевина, гуанидинхлорид и детергенты. Солюбилизированный в этих жестких условиях рекомбинантный белок для восстановления нативной конформации требует ренатурации, в процессе обработки денатурирующими агентами белки могут изменять свою структуру, а при концентрировании из разбавленных растворов, в которых проводится ренатурация, часто вновь выпадают в осадок.

В чем же причина такого аномального поведения рекомбинантных эукариотических белков в бактериальных клетках? Для поддержания белков в растворимом состоянии в клетках эукариот реализуются три основных механизма: компартментализация продуктов трансляции, белок-белковые взаимодействия и посттрансляционные модификации. Образующиеся в результате трансляции мРНК рибосомами в эндоплазматическом ретикулуме полипептидные цепи не распределяются хаотически в цитоплазме эукариотических клеток, но последовательно переходят через ряд компартментов, где претерпевают посттрансляционные модификации. При этом внутренние условия отдельных компартментов могут существенно различаться. Так, молекулы инсулина накапливаются *in vivo* в секреторных гранулах, pH в которых составляет 4,5–5,5. В то же время pH цитоплазмы бактериальных клеток приближается к 7,8, в этих условиях инсулин лишь слабо растворим. Многие гидрофобные эукариотические белки не существуют в растворимой форме в отсутствие фосфолипидов мембран. Белки молока являются интегральной частью мицелл, стабилизированных ионами Ca^{2+} .

Посттрансляционные модификации белков, в частности фосфорилирование, гидроксилирование, гликозилирование и ограниченный протеолиз, происходящие в определенных компартментах эукариотических клеток, также оказывают большое влияние на растворимость белков, их стабильность и биологические функции. Это особенно относится к секретируемым белкам эукариот. Поскольку в бактериальных клетках соответствующие системы посттрансляционных модификаций белков отсутствуют, в них не могут быть получены биологически полноценные рекомбинантные эукариотические полипептиды, что накладывает ограничения на использование бактерий в биотехнологии.

Наконец, третьим ключевым фактором, оказывающим влияние на растворимость белков, является их нативная конформация в растворе. Правильное сворачивание (фолдинг) полипептидных цепей некоторых белков в клетках эукариот обеспечивается специфическими белками, называемыми *шаперонами*, которые необходимы для эффективного формирования третичной структуры полипептидных цепей других белков, но не входят в состав конечной белковой структуры. Шапероны оказывают влияние лишь на конформацию полипептидных цепей и не действуют на ковалентные связи белковых молекул. К этому классу полипептидов относятся цитозольные белки теплового шока Hsp70 и Hsp60, а также гомологичные им белки, например белок, связывающийся с тяжелой цепью иммуноглобулинов (BiP). Шапероном является также ядерный нуклеоплазмин, обеспечивающий сборку нуклеосом. У *E. coli* соответствующие функции выполняют белки SecB, триггерный фактор, а также GroEL и GroES, обеспечивающие экспорт полипептидов из цитоплазмы и участвующие в морфогенезе бактериофагов. Два последних белка – GroEL и GroES, известные также как Cpn60 и Cpn10, кодируются генами оперона *groE*, регулируемого тепловым шоком. Белки, эволюционно родственные белку GroEL, получили название *шаперонинов*. Один из таких белков, обнаруженный в хлоропластах высших растений, участвует в сборке D-рибулозо-1,5-дифосфаткарбоксилазы-оксигеназы (Rubisco), которую часто используют в качестве модельного белка в исследованиях механизмов сворачивания полипептидных цепей.

Из других белков, участвующих в сворачивании полипептидных цепей, следует упомянуть так называемые фолдазы и изомеразы. Наиболее хорошо

изученными ферментами этой группы являются тиоредоксин и глутаредоксин, которые наряду с глутатионзависимым превращением рибонуклеотидов в дезоксирибонуклеотиды способны восстанавливать дисульфидные связи в белках, способствуя формированию у них правильной третичной структуры. Те же функции выполняет и дисульфидизомераза белков, осуществляя образование дисульфидных связей в секретируемых белках после их переноса через мембраны эукариотических клеток. Лимитирующей стадией в образовании правильной третичной структуры полипептидных цепей является *цис-транс*-изомеризация остатков пролина в полипептидных цепях. Этот процесс в клетках эукариот осуществляется с участием пролил-*цис-транс*-изомеразы. Поскольку все вышеперечисленные условия, необходимые для поддержания эукариотических рекомбинантных белков в растворимой форме, не могут быть реализованы в бактериальных клетках, в них происходит агрегация рекомбинантных белков с образованием телец включения. В настоящее время не существует универсального метода, который бы позволял получать рекомбинантные эукариотические белки в бактериальных клетках в нативном виде. В некоторых случаях уже понижение температуры выращивания рекомбинантных бактерий до 30° приводит к образованию полностью растворимых рекомбинантных белков. Это наблюдали в случае рекомбинантного γ -интерферона, А-цепи рицина, щелочного фактора роста фибробластов и в ряде других случаев.

Более универсальным подходом к получению рекомбинантных белков в растворимой форме является обеспечение секретируемого характера их экспрессии, т.е. секреции синтезирующихся рекомбинантных белков в питательную среду. Системы секреции у грамположительных и грамотрицательных бактерий интенсивно изучаются. Получение рекомбинантных эукариотических белков в секретируемой форме высокотехнологично, поскольку значительно облегчает их очистку. Кроме того, у секретируемых белков больше шансов приобрести нативную конформацию, так как сворачивание полипептидных цепей происходит в большом объеме, а следовательно, и более низкой их концентрации, что предотвращает агрегацию полипептидных цепей, не полностью сформировавших свою третичную структуру. Примером эффективного использования секреторной системы *E. coli* является получение рекомбинантного инсулино-подобного фактора роста I

(IGF-I) в секретируемой форме. В этой системе удавалось нарабатывать до 1 г рекомбинантного белка в 1 л бактериальной культуры.

Еще одним возможным, хотя и более труднореализуемым подходом к эффективному синтезу растворимых рекомбинантных белков в бактериальных клетках является введение экспрессирующихся генов шаперонов в клетки бактерий-хозяев. Так, сверхэкспрессия гомологичного *gro*-оперона в клетках *E. coli*, содержащих экспрессирующиеся гены большой и малой субъединиц Rubisco цианобактерии *Anacystis nidulans*, приводила к 5–10-кратному возрастанию внутриклеточного содержания нативного рекомбинантного белка Rubisco. В некоторых случаях внутриклеточную агрегацию сверхэкспрессирующихся белков удается предотвратить искусственной заменой отдельных аминокислот в полипептидных цепях методами белковой инженерии. Например, удаление семи аминокислот перед спиралью G в N-концевой части полипептидной цепи фрагмента Кленова ДНК-полимеразы I *E. coli*, а также замена нескольких гидрофобных аминокислот, боковые цепи которых экспонированы на поверхности спирали G, на остатки аспарагина полностью предотвращали внутриклеточную агрегацию фрагмента и образование агрегатов в процессе его очистки. В нативной молекуле ДНК-полимеразы I участок полипептидной цепи, в котором проведены замены, закрыт доменом 5'→3'-экзонуклеазы, отсутствующим у фрагмента Кленова, и не оказывает влияния на агрегационные свойства фермента.

Несмотря на определенные успехи в получении нативных рекомбинантных эукариотических белков в бактериальных клетках, эти системы экспрессии эукариотических генов в настоящее время еще далеки от совершенства. По-прежнему остается открытым вопрос о посттрансляционных модификациях рекомбинантных белков в бактериях, а для многих рекомбинантных белков не решена даже проблема их внутриклеточной растворимости. С этой точки зрения идеальными системами экспрессии рекомбинантных эукариотических генов являются гомологичные генетические системы, т.е. сами эукариотические клетки животных или растений (см. раздел 7.4). Действительно, такие системы экспрессии, хотя и более дорогие, с успехом используются в биотехнологии. При этом одним из ключевых факторов, обеспечивающих эффективное функционирование подобных систем, могут быть эукариотические экспрессирующие векторы, основные особенности

которых будут рассмотрены в следующем разделе.

7.2.7. Векторы для переноса ДНК в клетки животных и растений

Все основные принципы, используемые при конструировании бактериальных векторов, применимы и для получения векторов эукариотических клеток. Как и в случае бактерий, эукариотический вектор представляет собой небольшую молекулу ДНК, способную автономно реплицироваться в клетках животных. Помимо последовательностей нуклеотидов, обеспечивающих репликацию, эукариотические векторы могут содержать гены, используемые в качестве селективируемых маркеров, а также один или несколько уникальных сайтов рестрикции, по которым производится встраивание клонируемых последовательностей нуклеотидов ДНК. Поскольку непосредственное клонирование рекомбинантных ДНК в клетках животных было бы дорогостоящей и малоэффективной процедурой, эукариотические векторы используют, как правило, для получения экспрессии уже клонированных последовательностей нуклеотидов в клетках животных и растений, а сам процесс клонирования проводят в бактериях. Следовательно, эукариотические векторы, помимо всего прочего, должны быть челночными векторами. Для экспрессии в клетках рекомбинантные ДНК помещают под контроль регуляторных элементов, узнаваемых и используемых ферментативными системами эукариотических клеток.

В отличие от бактерий в клетках эукариот внехромосомные, автономно реплицирующиеся генетические элементы типа бактериальных плазмид встречаются редко. Поэтому при конструировании векторов, способных реплицироваться и осуществлять экспрессию в клетках животных, чаще всего используют регуляторные генетические элементы вирусов животных или, в зависимости от задач, эукариотических генов домашнего хозяйства, а также генов, для которых характерна тканеспецифическая экспрессия.

Экспрессирующий вектор pKSV-10. Структура теоретически возможного эукариотического вектора для экспрессии клонированных генов в клетках животных представлена на рис. II.12, из которого видно, что эукариотический вектор сохраняет все основные генетические элементы бактериальных векторов. Это, прежде всего, репликатор (область начала репликации *ori*), распознаваемый репликативными системами эукариотической

клетки. В роли репликатора чаще всего используют соответствующие последовательности нуклеотидов вирусов животных (например вируса SV40 или вируса Эпштейна–Барр). Особенностью функционирования такого репликатора является потребность в вирусных белках для инициации репликации. Так, в случае SV40 – это Т-антиген, тогда как у вируса Эпштейна–Барр – белок EBNA (Epstein-Barr nuclear antigen). Для того чтобы репликация векторов, в которых использованы вирусные репликаторы, осуществлялась без вируса-помощника, получили специальные линии клеток, которые стабильно продуцируют соответствующие вирусные белки. Например, в клетках линии COS синтезируется Т-антиген вируса SV40, а клетки линии Нер-EBNA-2 экспрессируют ген *EBNA*.

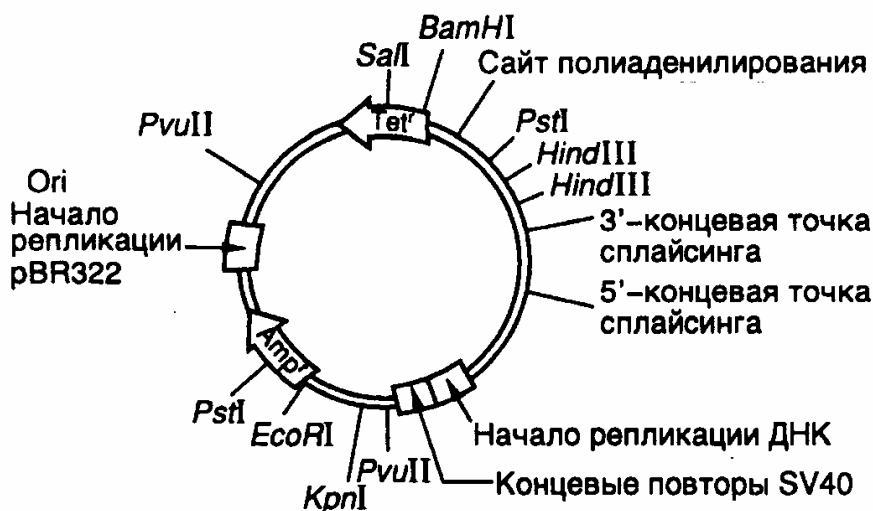


Рис. II.12. Эукариотический вектор рКСV-10

Часть бактериальной плазмиды, содержащей область начала репликации *ori*, соединена с транскрипционной единицей вируса SV40, 5'-концевая часть которой включает ранние промотор и репликатор вируса SV40, а 3'-конец содержит сайт полиаденилирования и интрон большого Т-антигена. Последовательности чужеродной ДНК, которые необходимо экспрессировать в клетках животных, клонируют по *Bgl*II-сайту рестрикции. Кроме того, в векторе имеются уникальные сайты для рестриктаз *Kpn*I, *Bam*HI, *Sal*I и *Eco*RI. Вектор может реплицироваться в клетках *E. coli* и в клетках обезьян линии COS.

В ряде случаев гены, необходимые для репликации эукариотического вектора, вводят непосредственно в вектор. Так же как и бактериальный, эукариотический вектор должен содержать один или несколько уникальных

сайтов рестрикции для клонирования требуемой последовательности нуклеотидов. В экспрессирующем векторе полилинкер располагают непосредственно за генетическими элементами, регулирующими транскрипцию (наибольшее значение из них имеют промотор и энхансер), перед сайтом полиаденилирования и терминатором транскрипции. Структура других кассет генов, обеспечивающих экспрессию рекомбинантных генов в клетках животных, будет рассмотрена в разделе 7.4.6.

Векторы растений. Специфика векторов для экспрессии рекомбинантных генов в клетках растений обусловлена главным образом регуляторными последовательностями нуклеотидов, которые обеспечивают эффективную транскрипцию генов в растительных тканях. Например, в известном векторе pCaMV CAT используются промотор и сайт полиаденилирования генов вируса мозаики цветной капусты, под контролем которых находится ген бактериальной хлорамфениколацетилтрансферазы (*cat*). Этот ген-репортер применяется для оптимизации условий введения векторной ДНК в клетки растений и может быть заменен на любой другой ген. Вектор pCaMV CAT является челночным, т.е. может реплицироваться как в клетках *E. coli*, так и растений. В первом случае селективируемым маркером служит, как обычно, ген устойчивости к ампициллину.

Вектор pCaMV CAT принадлежит к большой группе так называемых DMGT-векторов (DNA-mediated gene transfer), осуществляющих опосредованный ДНК перенос генов в клетки растений. Это означает, что в данном случае перенос генов в клетки растений опосредован ДНК вектора, так как они находятся в составе векторной молекулы. Для получения стабильных линий растительных клеток, которые бы содержали рекомбинантные ДНК в интегрированном в хромосомы состоянии и активно экспрессировали его, часто в состав DMGT-векторов вводят дополнительные селективируемые маркеры, которые придают клеткам, включившим экзогенную ДНК, определенный, легко обнаруживаемый фенотип. Такими маркерами могут быть, например гены неомизинфосфотрансферазы II (NPT II), гигромицинфосфотрансферазы (HPT), дигидрофолатредуктазы, в результате экспрессии которых клетки приобретают устойчивость к канамицину, гигромицину В или метотрексату соответственно.

7.3. Клонотеки генов

Любой индивидуальный ген занимает лишь небольшую часть генома живого организма. В то же время размер генома даже наиболее просто организованных бактерий в среднем составляет $2 \cdot 10^6$ п.о., а суммарный размер молекул ДНК, составляющих гаплоидный геном человека, по крайней мере, на три порядка больше. Из этого следует, что уникальные гены, представленные в гаплоидном геноме только одной копией, затеряны среди других последовательностей генома, и для работы с индивидуальными рекомбинантными ДНК требуется очистка от ненужного генетического материала. Такая задача в генной инженерии решается через создание *репрезентативных* (представительных) *клонотек* последовательностей нуклеотидов ДНК или, иначе говоря, *клонотек генов*.

7.3.1. Получение клонотек генов

Клонотека генов представляет собой набор разных последовательностей нуклеотидов ДНК, клонированных в составе векторных молекул, которые в сумме составляют весь геном исследуемого организма или какую-либо известную его часть. При этом репрезентативная клонотека должна заключать в себе с высокой долей вероятности любую последовательность нуклеотидов изучаемого генома. Как уже подробно обсуждалось выше, для большинства генов эукариот характерна интрон-экзонная структура, а интроны, присутствующие в первичных транскриптах таких генов, вырезаются в процессе сплайсинга с образованием зрелых молекул мРНК. Для получения экспрессии эукариотических генов в клетках прокариот, а также изучения тканеспецифической экспрессии генов возникает необходимость в получении кодирующих последовательностей эукариотических генов без интронов. В этом случае задача решается через создание репрезентативных клонотек кДНК, в которых с высокой вероятностью представлены последовательности нуклеотидов мРНК, синтезирующихся в тканях, культурах тканей или отдельных соматических клетках.

На рис. II.8 представлена схема конструирования клонотеки геномной ДНК с использованием космидного вектора. Следует иметь в виду, что общие принципы получения клонотек генов приложимы и к любым другим векторам.

На первом этапе конструирования осуществляют подготовку вектора и клонируемой ДНК. Космидный вектор расщепляют рестриктазами *Bam*HI и *Sma*I, причем рестриктаза *Sma*I расщепляет ДНК с образованием "тупых" концов. В результате образуются два "плеча" космидного вектора, содержащих область начала репликации *ori*, используемую системой репликации бактериальных клеток, и два селектируемых маркера, которые представляют собой гены устойчивости к антибиотикам – ампициллину (Amp^r) и канамицину (Can^r), а также два *cos*-сайта хромосомы бактериофага λ . Параллельно со всеми необходимыми предосторожностями получают препараты высокомолекулярной ДНК, которую необходимо клонировать.

Геномную ДНК подвергают частичному гидролизу мелкощепящей рестриктазой *Mbo*I (узнает последовательность из четырех нуклеотидов GATC, которые комплементарны "липким" концам, образуемым рестриктазой *Bam*HI). Время рестрикции и концентрацию фермента подбирают таким образом, чтобы средняя длина образующихся фрагментов ДНК составляла 35–45 т.п.о. Обогащенную фракцию фрагментов ДНК подобного размера далее получают центрифугированием смеси рестрикционных фрагментов в градиенте концентрации сахарозы и лигируют с приготовленными "плечами" вектора в условиях, при которых образование сшивок по "тупым" концам минимально. При этом, помимо требуемых молекул рекомбинантных ДНК, могут образовываться и артефактные молекулы, которые не будут упаковываться в фаговые частицы либо из-за их неоптимального размера, либо вследствие неправильной ориентации *cos*-сайтов.

Полученную после лигирования сложную смесь молекул рекомбинантных ДНК упаковывают в фаговые частицы стандартным способом, и образовавшимися фаговыми частицами заражают подходящие бактериальные клетки. В итоге колонии бактериальных клеток, которые выросли в присутствии антибиотиков, должны заключать в себе различные последовательности нуклеотидов клонируемой ДНК приблизительно одинаковой длины.

На этом частном примере был рассмотрен один из подходов к конструированию клонотек геномной ДНК с использованием космидного вектора, который иллюстрирует все основные этапы получения клонотек генов с применением и других векторов, плазмидных или фаговых. Во всех случаях для получения репрезентативных клонотек генов следует помнить о

необходимости случайной фрагментации клонируемой высокомолекулярной ДНК, с тем чтобы все участки генома в образующейся смеси фрагментов были представлены равновероятно. Кроме того, размеры образуемых фрагментов ДНК должны соответствовать емкости вектора, используемого для их клонирования. Так, для клонирования в космидах необходимо получать фрагменты ДНК длиной 35–45 т.п.о., тогда как для клонирования в фаговых векторах типа Charon и плазмидах эти размеры должны составлять соответственно 20–24 и 1,5–3,0 т.п.о.

Чем короче фрагменты, используемые для получения библиотек генов, и чем сложнее исследуемый геном, тем большее число клонов необходимо получить, чтобы клонотека была полной. Например, для того чтобы создать геномную клонотеку бактерий из фрагментов ДНК длиной в 20 т.п.о., в которой любая последовательность была бы представлена с вероятностью 99%, в ней необходимо иметь 460 клонов, тогда как для млекопитающих это число возрастает до 600000, а для покрытосеменных растений оно еще в ~10 раз больше. Отсюда следует, что в случае получения репрезентативных клонотек генов бактерий для решения такой задачи вполне достаточно плазмидных векторов, несмотря на их небольшую емкость. В то же время использование плазмид для создания полных клонотек генов млекопитающих или растений совершенно бесперспективно, так как потребовало бы поддержания в клонотеке огромного количества клонов, многие из которых с большой вероятностью могут быть утрачены.

Выше были обсуждены условия, которые необходимо соблюдать для получения репрезентативных клонотек геномной ДНК. Однако во многих случаях не возникает необходимости работать с полными клонотеками генов, и исследователи ограничиваются неполными клонотеками, обогащенными отдельными участками генома. Так, клонотеки кДНК заключают в себе последовательности нуклеотидов, которые в виде зрелых мРНК специфически экспрессируются в различных клетках или тканях. То же можно сказать и о клонотеках повторяющихся последовательностей нуклеотидов, отобранных из препарата суммарной ДНК на основании кинетики их реассоциации.

7.3.2. Введение рекомбинантных ДНК в клетки

После завершения лигирования векторных молекул с клонируемыми фрагментами ДНК реакционную смесь с продуктами реакции уже можно рассматривать как библиотеку генов. Полученные таким образом рекомбинантные молекулы ДНК хранятся длительное время в замороженном состоянии, и по мере необходимости из них отбирают аликвоты реакционной смеси, содержащие требуемое количество рекомбинантных молекул, и вводят в клетки микроорганизмов для клонирования и дальнейшей работы.

Процесс поглощения экзогенной ДНК бактериальными клетками, сопровождающийся приобретением ими новых генетических маркеров, называют *генетической трансформацией*. В том случае, если бактериальные клетки поглощают ДНК бактериофагов и в результате происходит образование зрелых фаговых частиц, говорят о *трансфекции* фаговой ДНК. Поглощение экзогенной ДНК в процессе трансформации и трансфекции может осуществляться лишь *компетентными клетками*. Под компетентностью понимается такое состояние бактериальных клеток, при котором они способны сорбировать экзогенную ДНК на своей поверхности и поглощать ее, после чего экзогенная ДНК может быть интегрирована в бактериальную хромосому или существовать в виде внехромосомного элемента (эписомы, плазмиды).

Для многих грамположительных бактерий компетентность является естественным свойством, которое становится максимально выраженным на определенных этапах роста культуры. В то же время у наиболее важных для генной инженерии грамотрицательных клеток *E. coli* естественная компетентность вообще отсутствует и клетки приобретают ее лишь в искусственных условиях. В настоящее время установлено, что при низких температурах и в присутствии двухвалентных катионов клетки *E. coli* и экзогенная ДНК продуктивно взаимодействуют друг с другом и экзогенная ДНК может с высокой эффективностью проникать внутрь бактериальных клеток. Частоту трансформации и трансфекции клеток *E. coli* можно повысить разными способами. Среди них наиболее популярными являются тепловой шок, введение в инкубационную смесь одновалентных катионов (Rb^+), хлорида гексаминкобальта(III) или выращивание бактериальных клеток на среде с повышенным содержанием ионов Mg^{2+} (10–20 мМ).

Современные методики трансформации бактериальных клеток позволяют получать до 10^7 – 10^8 трансформантов или трансфектантов на 1 мкг суперскрученной плазмидной ДНК. Поскольку трансформация и трансфекция проводятся при низких концентрациях экзогенной ДНК, а для появления у трансформированной бактерии устойчивости к антибиотикам или развития фаговой бляшки достаточно проникновения внутрь бактериальной клетки одной молекулы рекомбинантной ДНК, образовавшиеся колонии или бляшки трансформантов и трансфектантов содержат, как правило, только один тип рекомбинантных молекул ДНК с идентичными вставками клонируемых последовательностей нуклеотидов.

Еще более высокой эффективности трансформации (до 10^9 – 10^{10} трансформантов на 1 мкг ДНК) достигают при использовании *электропорации*. В этом случае клетки вместе с плазмидной или фаговой ДНК помещают в небольшую ячейку между двумя электродами и подвергают кратковременному (10–100 мкс) воздействию электрического поля высокой напряженности. Быстрая поляризация клеточных мембран приводит к их обратимому повреждению (образованию пор), сопровождаемому повышением их проводимости и проницаемости для макромолекул. В это время происходит эффективный перенос экзогенных молекул ДНК внутрь клеток, подвергнутых электрическому шоку. После того как рекомбинантные молекулы ДНК введены в бактериальные клетки и получены бактериальные колонии или фаговые бляшки, каждая из которых содержит рекомбинантные ДНК только одного типа, возникает важная задача отбора тех колоний или бляшек, которые заключают в себе искомые последовательности нуклеотидов рекомбинантных ДНК.

7.3.3. Методы скрининга клонотек генов

Все методы получения из клонотек генов требуемых последовательностей нуклеотидов можно разделить на две группы. При использовании первой группы методов рекомбинантные бактерии или фаговые частицы исследуют на наличие в них искомым последовательностей нуклеотидов. Во втором случае присутствие этих последовательностей обнаруживают косвенно, по появлению в бактериальных клетках или фаговых лизатах продуктов экспрессии искомым генов – РНК, белков или ферментативной активности.

Для выявления конкретных последовательностей нуклеотидов в клонотеках генов рекомбинантные бактерии или фаги рассеивают на чашках Петри и образовавшиеся колонии или бляшки переносят на нитроцеллюлозные или нейлоновые фильтры. Для этого стерильные фильтры накладывают на поверхность чашек, в результате чего часть бактериальных клеток или фаговых частиц прочно связывается с поверхностью фильтров. Расположение их на фильтрах и поверхности чашек Петри точно совпадает. Бактериальные клетки или фаговые частицы, сорбированные на фильтрах, лизируют щелочью, что сопровождается одновременной денатурацией заключенной в них ДНК. Освободившуюся ДНК фиксируют на фильтрах и гибридизуют с меченым олигонуклеотидным зондом, последовательность нуклеотидов которого точно соответствует последовательности нуклеотидов искомой ДНК. В результате гибридизации происходит специфическое прочное связывание зонда с идентичными и/или гомологичными последовательностями нуклеотидов в ДНК, содержащейся в отдельных бактериальных колониях или бляшках.

После отмывки фильтров от несвязавшегося радиоактивного зонда их высушивают и прикладывают к рентгеновской пленке. Образование специфических гибридов обнаруживают после проявления рентгеновской пленки по появлению на ней четких гибридизационных сигналов в виде темных пятен, положение которых точно соответствует таковому определенных колоний или бляшек на исходной чашке Петри. В результате проведения такого исследования точно идентифицируются бактериальные колонии или фаговые бляшки, содержащие искомые рекомбинантные ДНК. С этого момента с помощью обычных микробиологических и биохимических методов можно получать неограниченное количество идентичных копий клонированной последовательности нуклеотидов ДНК для дальнейших исследований. В ряде случаев, когда нет необходимости в скрининге большого числа клонов, гибридизацию с зондами можно заменить ПЦР. При таком подходе в качестве источника матричной ДНК используют суспензию бактериальных клеток или фаговых частиц отдельных клонов без специальной очистки матричных нуклеиновых кислот.

Часто бывает, что первичная структура искомого гена неизвестна. Тогда применяют один из следующих способов. Во-первых, можно очистить небольшое количество белка, кодируемого клонируемым геном, и определить

последовательность его нескольких N-концевых аминокислот. На основании этих данных с учетом вырожденности генетического кода и частоты встречаемости аминокислотных остатков у исследуемого организма синтезируют олигонуклеотидные зонды, которые также используют для гибридизации *in situ*, аналогично тому как было описано выше. Во-вторых, клонотеку генов или кДНК получают с использованием экспрессирующих векторов. В этом случае если 5'-концевая часть клонируемого гена будет находиться близко от промотора вектора или его кодирующая часть будет соединена в одной рамке считывания с иницирующим ATG-кодоном, то в бактериальных клетках или фаговых лизатах можно обнаружить появление специфического белкового продукта или ферментативной активности. Если рекомбинантный белок неактивен и представляет собой лишь часть полипептидной цепи нативного фермента, его обнаруживают иммунохимическими методами. В ряде случаев отбор нужных клонов может проводиться с помощью генетических методов с использованием специально сконструированных векторных плазмид.

Рассмотренные способы не исчерпывают всех методических подходов, используемых для идентификации определенных последовательностей нуклеотидов в клонотеках генов. Но и эти часто применяемые методы дают представление о том, каким путем сегодня можно выделить индивидуальный ген и добиться его экспрессии в новом генетическом окружении.

7.4. Эукариотические системы экспрессии рекомбинантных генов, основанные на культурах клеток

Выделение любого нового рекомбинантного гена описанными выше методами неизбежно заканчивается попытками получения его полноценной экспрессии в искусственных генетических системах. Только на первый взгляд экспрессия рекомбинантных генов в бактериальных клетках под контролем хорошо изученных регуляторных элементов бактерий представляется наиболее простой задачей. На практике, как уже обсуждалось ранее, экспрессия эукариотических генов в бактериях происходит крайне неэффективно из-за образования нерастворимых телец включения и отсутствия необходимых посттрансляционных модификаций рекомбинантных полипептидных цепей. Для преодоления этих затруднений в последнее время широко используются

культивируемые клетки животных и растений, в геном которых рекомбинантные гены вводятся с помощью трансфекции.

Однако эффективный синтез рекомбинантных белков зависит не только от используемых клеточных линий. Основное влияние на этот процесс оказывают конструкция экспрессирующего вектора, а также сам метод введения рекомбинантных ДНК в эукариотические клетки. В последнем случае низкая эффективность доставки рекомбинантных молекул в клетки и неоптимальная локализация сайтов их интеграции в геном могут свести на нет все достоинства клеточных линий и экспрессирующих векторов. Для повышения уровня экспрессии рекомбинантных генов, стабильно интегрированных в геном культивируемых клеток, разработаны методы их селективной амплификации. С использованием одного из таких подходов получены мутантные сублинии клеток яичников китайских хомячков СНО, в которых амплификация происходит в присутствии селектирующего агента, например метотрексата. Другие системы амплификации основаны на подавлении экспрессии жизненно важных ферментов, например глутаминсинтетазы или аденозиндезаминазы, под действием специфических ингибиторов: метионинсульфоксимины или 2'-дезоксикоформина соответственно. Механизмы амплификации генов под действием ингибиторов ферментов будут подробнее рассмотрены ниже на примере метотрексата (см. раздел 7.4.1). Клеточные линии и клоны, стабильно продуцирующие рекомбинантные белки, как правило, получают путем отбора из пула гетерогенных клеток. И, наконец, еще одним существенным условием успешной экспрессии рекомбинантных генов является оптимизация самого процесса культивирования клеток, включая тщательный подбор культуральных сред.

Для получения рекомбинантных белков кроме вышеупомянутого метода стабильно трансфецированных линий клеток используют еще два современных подхода. С целью наработки небольшого количества рекомбинантных белков и контроля функциональной целостности генно-инженерных конструкций часто применяют временно экспрессирующиеся векторы, которые обладают способностью к репликации в культивируемых клетках COS во внехромосомном состоянии. В этом случае источником рекомбинантных белков являются супернатанты клеток. При другом подходе рекомбинантные белки синтезируют в культивируемых клетках насекомых после их заражения векторами на основе

генома бакуловирусов. Последний подход становится все более популярным, так как допускает стабильную экспрессию многих рекомбинантных генов, продукты которых могут находиться как внутри клеток, так и в ассоциированном с поверхностью клеток состоянии, или они могут секретироваться в культуральную жидкость.

Недавно швейцарскими исследователями (С. Гайссе с соавт., 1996 г.) было проведено сравнительное изучение эффективности экспрессии гена цитокина человека – фактора, ингибирующего лейкозы (leukemia inhibitory factor – huLIF) в разных эукариотических системах. Фактор huLIF представляет собой белок среднего размера, полипептидная цепь которого содержит несколько потенциальных сайтов N-гликозилирования и в процессе фолдинга образует три дисульфидные связи. Поскольку полученные результаты представляют большой практический интерес для генной инженерии и биотехнологии, так как иллюстрируют многие основные принципы функционирования рекомбинантных генов в гетерологичных системах экспрессии, они будут кратко рассмотрены ниже.

7.4.1. Клетки яичников китайских хомячков (линия CHO)

Эта линия клеток и ее многочисленные производные часто используются для синтеза рекомбинантных белков после предварительной эндогенной амплификации соответствующих рекомбинантных генов, введенных в клетки с помощью трансфекции. Спонтанная амплификация генов в нормальных клетках животных является редким событием, а в трансформированных (опухолевых) клетках она происходит с частотой 10^{-4} – 10^{-6} за клеточную генерацию. Частота амплификации определенных генов может быть повышена в результате некоторых экзогенных воздействий, включая инкубацию клеток с гидроксимочевинной, афидиколином, канцерогенами, а также под действием ультрафиолетового света или γ -излучения и в условиях гипоксии. Амплификация гена обычно начинается с его вырезания из хромосомы в составе протяженного сегмента ДНК, сопровождаемого образованием кольцевой структуры, которая продолжает существовать в виде внехромосомного элемента (*двойной минихромосомы* (double-minutes – DM)), способного к автономной репликации. Предполагают, что DM является

промежуточным продуктом амплификации, образующимся в условиях низкого давления селектирующего фактора окружающей среды. На более поздних стадиях амплифицирующаяся ДНК после повторной интеграции в геном обнаруживается преимущественно в составе *удлиненного участка хромосомы* (extended chromosomal region – ECR), иначе называемого *гомогенно окрашивающейся областью* (homogeneously staining region – HSR). На этой стадии амплифицированный участок генома клеток, а следовательно, и их фенотип становятся более стабильными.

Для биотехнологических целей в системе амплификации рекомбинантных генов в качестве селектируемого маркера часто используют ген дигидрофолатредуктазы (ДГФР), интегрированный в экспрессирующий вектор. ДГФР катализирует превращение фолиевой кислоты в тетрагидрофолат, необходимый для синтеза эукариотическими клетками глицина, тимидинмонофосфата и пуриновых оснований. В этой связи клетки СНО, в которых ген ДГФР инактивирован под действием мутации, не растут на средах без нуклеозидов и приобретают эту способность после трансфекции геном ДГФР. Растущие трансфектанты далее отбираются по признаку амплификации гена ДГФР на фоне увеличивающихся концентраций метотрексата – ингибитора ДГФР, в питательной среде, так как увеличение числа копий гена будет придавать клеткам устойчивость к ингибитору в больших концентрациях. После проведения множественных раундов селекции получают популяцию клеток, содержащих до нескольких сотен копий гена ДГФР.

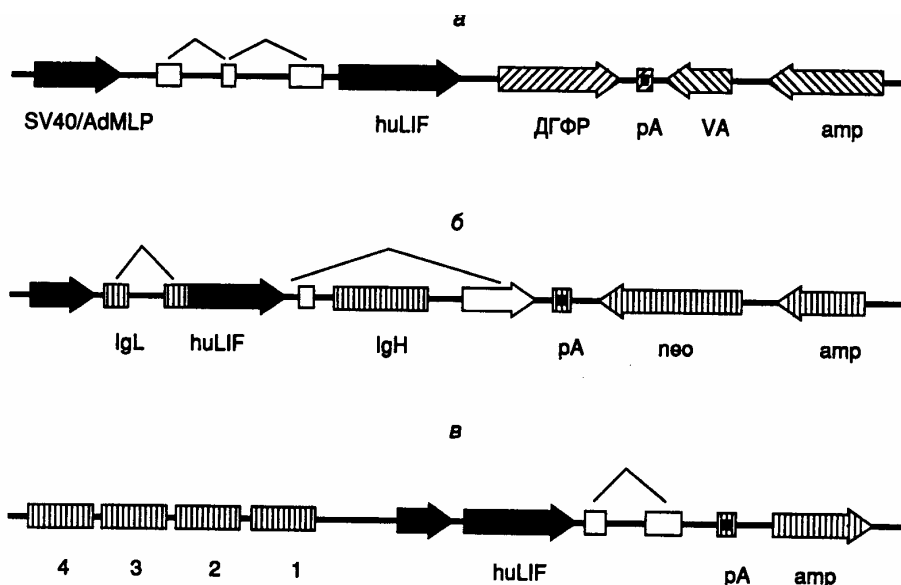


Рис. II.13. Кассеты генов в экспрессирующих векторах, использованные для получения фактора, ингибирующего лейкозы, в культурах клеток

а – кДНК huLIF (черные стрелки), включающая лидерные последовательности (незаштрихованные прямоугольники), была интегрирована в экспрессирующий вектор рХМТ3 под контроль гибридного промотора, расположенного между точкой начала репликации вируса SV40 и основным поздним промотором аденовируса (SV40/AdMLP), ДГФР, amp, neo – гены дигидрофолатредуктазы и устойчивости к ампициллину и неомицину соответственно, VA – трансляционный энхансер; pA – сайт полиаденилирования;

б – вектор рSV2neo-Ig-huLIF, сконструированный на основе транскрипционных и трансляционных регуляторных элементов генов тяжелой (IgH) и легкой (IgL) цепей иммуноглобулинов. Транскрипция кассеты контролируется мышинным промотором гена IgL κ и энхансером гена IgH. Транскрипт, содержащий последовательность huLIF, в результате сплайсинга объединяется с экзоном 2 гена IgL, кодирующего лидерную последовательность, которая обеспечивает секрецию зрелого huLIF клетками;

в – вектор рGES-LIF, прямоугольники 1–4 – регуляторная область кластера глобиновых генов, короткая черная стрелка – промотор β-глобинового гена, ломаной линией отмечены последовательности, удаляемые сплайсингом. Во всех случаях стрелки указывают направление транскрипции

Размер амплифицированного геномного локуса значительно больше размера селектируемого гена и может составлять несколько сотен тысяч пар оснований. Трансфекция клеток геном ДГФР и исследуемым рекомбинантным геном, заключенными в один экспрессирующий вектор, сопровождается их

интеграцией в геном, будет приводить к их совместной амплификации на фоне метотрексата и усилению синтеза требуемого рекомбинантного белка, как следствие увеличения дозы его гена. Данные о возможности сохранения высокого уровня экспрессии гена в отсутствие селектирующего агента, т.е. генетической стабильности продуцентов, противоречивы. Результат зависит от генотипа конкретных клонов клеток, в том числе и локализации сайта интеграции экспрессирующего вектора в геном клеток СНО. Эффективным экспрессирующим вектором, используемым в описываемой системе, является двухцистронная молекула ДНК, в которой исследуемый рекомбинантный ген находится под контролем сильного промотора перед геном ДГФР, что обеспечивает высокое сопряжение синтеза рекомбинантного белка с уровнем устойчивости клеток к метотрексату (рис. II.13,а).

7.4.2. Клетки мышинной миеломы (линия Sp2/0)

Клетки миеломы находят широкое применение для получения *гибридом* – линий клеток, производящих моноклональные антитела определенной специфичности. Поскольку эти клетки осуществляют эффективную секрецию рекомбинантных белков, они хорошо изучены в отношении экспрессии в них соответствующих рекомбинантных генов, введенных с помощью трансфекции. Кроме того, такие клетки способны расти в суспензионной культуре (без прикрепления к поверхности субстрата), что облегчает их культивирование при необходимости крупномасштабной наработки.

Линия клеток мышинной миеломы Sp2/0 Ag14, утративших способность синтезировать или секретировать иммуноглобулины, часто используется для производства рекомбинантных белков. Клетки этой линии легко трансфецируемы рекомбинантными ДНК и могут выращиваться в больших количествах на питательной среде без дорогостоящей сыворотки. Клетки Sp2/0 экспрессируют эндогенный ген ДГФР, что делает невозможным их прямое применение для амплификации рекомбинантных ДНК с использованием метотрексата. Это затруднение обычно преодолевают путем введения в экспрессирующий вектор мутантного гена ДГФР, в котором точковая мутация, сопровождаемая заменой Т→G, приводит к замещению Leu→Arg в положении 22 полипептидной цепи ДГФР, следствием чего является понижение сродства

фермента к метотрексату в 270 раз. Благодаря этому после трансфекции может быть проведен отбор клеток, содержащих мутантный ген ДГФР вместо нормального, путем постепенного увеличения концентрации метотрексата в среде. При умеренных концентрациях ингибитора эндогенный ген ДГФР полностью инактивируется, тогда как мутантный ген остается активным и может быть амплифицирован в результате описанного выше механизма после повышения концентрации метотрексата в среде. Однако уровень амплификации гена ДГФР и ассоциированного с ним рекомбинантного гена в такой системе как правило ниже, чем в обычной, так как меньшее число внутриклеточных копий мутантного гена требуется для обеспечения необходимого уровня устойчивости клеток к ингибитору. Кроме того, было показано, что высокий уровень внутриклеточной экспрессии рекомбинантных генов, введенных в клетки Sp2/0 с помощью трансфекции, может обеспечиваться после инкубации с метотрексатом за счет увеличения скорости их транскрипции, а не в результате их амплификации. Такой результат может быть получен в результате введения множественных копий гена ДГФР при трансфекции. Структура одного из векторов, пригодного для использования с миеломными клетками Sp2/0, представлена на рис. II.13,б). Транскрипция трансгенов в этом векторе обеспечивается промоторно–энхансерными элементами гена иммуноглобулина с участием эндогенных регуляторных факторов клеток миеломы.

7.4.3. Клетки селезенки мышей (линия MEL)

Обе системы экспрессии, описанные выше, базируются на амплификации трансгенов, обеспечивающей высокий уровень внутриклеточного синтеза кодируемых ими рекомбинантных белков в отобранных клонах клеток. Тем не менее, в природе уникальные гены (т.е. присутствующие в виде одной копии на гаплоидный геном) могут экспрессироваться с очень высокой эффективностью. Примером этого являются, например иммуноглобулины, секретируемые плазматическими клетками, или глобины, синтезируемые в больших количествах клетками эритроидного ряда. Высокий уровень транскрипции глобинового локуса становится возможным не только благодаря функционированию эффективных промотора и энхансера, но и в результате присутствия так называемой

доминантной регуляторной области (dominant control region – DCR), изменяющей пространственную структуру хроматина всего локуса. DCR расположена за ~50 т.п.о. перед кластером глобиновых генов и вначале была идентифицирована благодаря присутствию в ней нескольких сайтов, гиперчувствительных к ДНКазе. Путем комбинирования последовательностей нуклеотидов, окружающих сайты, удалось сконструировать небольшую искусственную DCR, обладающую всеми основными свойствами исходной. Рекомбинантные гены, объединенные в векторе с такой DCR, экспрессировались с высокой эффективностью в клетках эритроидного ряда независимо от места их интеграции в геном и не требовали для этого предварительной амплификации. Таким образом, рассматриваемая система позволяет избежать отрицательного влияния эффекта положения у рекомбинантных генов, что избавляет от необходимости скрининга большого числа трансфецированных клеток для получения высокопродуктивных продуцентов рекомбинантных белков.

Для биотехнологических целей в настоящее время часто используется линия клеток MEL-E9, которая является производной линии клеток MEL-NP5, полученных из селезенки мышей NFS, зараженных дефектным по репликации вирусом, образующим фокусы селезенки (spleen focus forming virus – SSFV). Эти клетки не образуют вирусных частиц, но вызывают возникновение опухолей после инъекции сингенным мышам. Они не претерпевают спонтанной дифференцировки, однако под действием экзогенных факторов (диметилсульфоксида, гемина или гексаметиленбисацетамида) дифференцируются в клетки, содержащие гемоглобин.

На рис. II.13,в изображена схема экспрессирующего вектора, применяемого с этой линией клеток. В нем область DCR соединена с промотором, экзоном 2, интроном 2 и экзоном 3 β -глобинового гена, поскольку показано, что присутствие последовательностей нуклеотидов интрона 2 абсолютно необходимо для эффективной экспрессии генов, направляемой этим промотором. Рекомбинантный ген может быть встроен справа или слева от интрона 2. После успешной трансфекции этот экспрессирующий вектор обеспечивает независимую от эффекта положения и числа копий тканеспецифическую экспрессию рекомбинантных генов в клетках эритроидного ряда.

7.4.4. Клетки африканской зеленой марышки (линия COS)

Получение временной экспрессии генов в клетках COS часто используется для быстрой наработки рекомбинантных белков и ДНК. При конструировании клеток COS клетки зеленой марышки CV-1 были трансформированы вирусом SV40 с дефектной областью начала репликации. В результате были получены три линии клеток (COS-1, 2 и 7), конститутивно экспрессирующих большой Т-антиген вируса. После проведения трансфекции клеток экспрессирующими плазмидами, содержащими область начала репликации вируса SV40, последняя эффективно взаимодействует с эндогенным Т-антигеном, что сопровождается внехромосомной репликацией плазмиды с образованием большого числа ее копий в цитоплазме клеток. Если рекомбинантный ген в такой плазмиде находится под контролем подходящего промотора, то это приводит к внутриклеточному образованию рекомбинантных белков в больших количествах.

Число копий плазмид, введенных в клетки COS, достигает максимума через ~48 ч после трансфекции. Далее внутриклеточный уровень плазмидной ДНК начинает постепенно снижаться из-за ее цитопатического действия и гибели клеток. В результате системы, основанные на клетках COS, не могут быть использованы для крупномасштабной наработки рекомбинантных белков в течение длительного промежутка времени. Максимальное внутриклеточное содержание рекомбинантных белков в этих системах наблюдают через 72 ч после трансфекции, и их синтез продолжается в течение последующих 5–10 дней на фоне медленного снижения количества клеток в культуре, что позволяет все же использовать клетки COS для препаративного синтеза рекомбинантных белков. С этой целью трансфекции одновременно подвергают до 10^8 клеток с последующим их выращиванием на роллерах или микроносителях, сопровождаемым многократным сбором культуральной жидкости. Такой подход позволяет получать до нескольких миллиграммов рекомбинантного белка из вышеописанного пула трансфецированных клеток.

Эффективность системы экспрессии, основанной на клетках COS, зависит, в первую очередь, от природы рекомбинантного белка, которая определяет его чувствительность к внутриклеточным протеиназам, а также метода трансфекции, соотношения числа клеток и молекул рекомбинантных

плазмид, используемых для трансфекции, и состава питательной среды. Пример экспрессирующей кассеты генов, используемой в этой системе, представлен на рис. II.13,а. Такой фрагмент ДНК может быть интегрирован в подходящую коммерческую плазмиду, в частности в экспрессирующий вектор рХМТ3.

7.4.5. Клетки насекомых, зараженные бакуловирусами

Многочисленное семейство бакуловирусов, размножающихся в клетках беспозвоночных, обладает геномом в виде двухцепочечной кольцевой ковалентно замкнутой ДНК длиной в 80–220 т.п.о. Круг хозяев бакуловирусов ограничен главным образом членистоногими. Клетки млекопитающих могут быть заражены бакуловирусами только при высокой множественности инфекции. При этом вирус проникает в клетки, но в них не воспроизводится, не экспрессирует рекомбинантные гены с промоторов насекомых и не способен находиться (персистировать) в клетках животных длительное время.

Интерес к бакуловирусам как потенциальным переносчикам рекомбинантных генов и системам, обеспечивающим их экспрессию, стал возникать после того, как была обнаружена их способность к синтезу большого количества вирусных белков на поздних стадиях инфекции. В частности, это касается вирусных полиэдрина и белка р10, гены которых находятся под контролем сильных поздних промоторов и которые могут быть удалены без ущерба для размножения вирусов. Поэтому замещение указанных генов рекомбинантными сопровождается усиленной экспрессией последних.

Первоначально использование бакуловирусов с культурами клеток вызывало большие затруднения из-за экзотичности объекта исследования. Однако с ростом интереса к этим системам экспрессии были разработаны многочисленные варианты рекомбинантных вирусов и векторов на их основе, включая челночные векторы для клеток дрожжей и бактерий. Кроме того, успех этой системы связан с получением удобных линий клеток насекомых. В настоящее время наиболее часто используемыми клеточными линиями остаются IPLB-Sf-21-AE (получена из *Spodoptera frugiperda*), ее клональная производная Sf9, а также BTI-TN-5B1–4 и TN368 из *Trichopsula ni*. Современные клеточные линии насекомых выращивают в виде суспензионных культур на роллерах и в ферментерах на питательных средах, не содержащих фетальной

сыворотки.

Для проявления в полной мере биологической активности многих рекомбинантных белков требуются посттрансляционные модификации их полипептидных цепей. Оказалось, что клетки насекомых обладают способностью производить большое число таких модификаций, включая гликозилирование, фосфорилирование, ацилирование остатками жирных кислот и амидирование. Рекомбинантные белки могут претерпевать в них адекватный протеолитический процессинг путем удаления сигнальных последовательностей аминокислот и переноситься в соответствующие клеточные компартменты. Были проведены интенсивные исследования механизмов N-гликозилирования в клетках насекомых, которые показали, что хотя оно и происходит, но не соответствует в полной мере таковому клеток млекопитающих.

7.4.6. Сравнение эффективности рассмотренных систем экспрессии

Проведено сравнение эффективности рассмотренных выше систем экспрессии эукариотических рекомбинантных генов с использованием гена huLIF в качестве модели. *In vivo* этот белок с молекулярной массой 32–67 кДа в зависимости от источника получения синтезируется под контролем уникального гена, содержащего три экзона. Фрагмент кДНК, кодирующий зрелый процессированный huLIF, клонировали в составе трех векторов, структура которых изображена на рис. II.13,а–в. Для трансфекции клеток COS и CHO использовали одну и ту же конструкцию (см. рис. II.13,а). Оказалось, что все пять исследованных линий клеток обладали способностью к синтезу рекомбинантного huLIF, обладающего биологической активностью. В клетках линий CHO, Sp2/0 и MEL после проведения электрофореза продукт обнаруживали в виде отчетливой, хотя и широкой полосы. В то же время рекомбинантный белок в клетках, зараженных бакуловирусами, был представлен дискретными фракциями меньшей молекулярной массы, что указывало на его неполное гликозилирование. Рекомбинантный материал в клетках COS был представлен дисперсной фракцией с молекулярной массой 20–40 кДа, что интерпретировалось в пользу ограниченной способности этих клеток осуществлять процессинг сильно гликозилированных полипептидных цепей.

Стабильные клетки-продуценты, созданные на основе амплификации рекомбинантных генов, позволяют получать рекомбинантные продукты в более высоких титрах по сравнению с клетками, адаптированными к "быстрой" кратковременной экспрессии рекомбинантных белков (клетки MEL или COS). Для получения высокого уровня экспрессии рекомбинантных белков во всех обсуждаемых системах необходимо, прежде всего, учитывать особенности структуры белка, который должен быть синтезирован, и конструировать экспрессирующие векторы в соответствии с требованиями, предъявляемыми теми или иными клетками. В ряде случаев предварительные исследования такого рода бывает удобнее провести в бесклеточных белоксинтезирующих системах, которые позволяют иногда получать даже препаративные количества рекомбинантных белков.

7.5. Бесклеточные белоксинтезирующие системы

Среди искусственных систем биосинтеза белка важное место занимают бесклеточные системы. Любая бесклеточная система создается, прежде всего, для моделирования конкретных биохимических процессов, происходящих в живом организме, и во время функционирования воспроизводит некоторые существенные особенности жизнедеятельности клетки. В генной инженерии бесклеточные белоксинтезирующие системы часто используются для исследования кодирующего потенциала и механизмов экспрессии клонированных генов *in vitro*, а также на промежуточных этапах конструирования рекомбинантных генов для идентификации мРНК или фрагментов ДНК по кодируемым белкам.

Большим преимуществом бесклеточных систем перед целыми клетками является доступность их отдельных компонентов для экспериментальных воздействий. Такие системы позволяют исследовать влияние различных экзогенных факторов на их функционирование (ионные условия, pH, ингибиторы и активаторы и т.п.). Кроме того, в бесклеточной системе можно легко заменять отдельные компоненты или непосредственно воздействовать на них в изолированном состоянии и затем по реакции системы познавать их функциональную значимость. Большинство результатов, полученных с помощью бесклеточных систем, невозможно было бы иметь при использовании живых клеток, так как последние при нарушении гомеостаза нередко гибнут.

При этом бывает трудно определить, какой же компонент оказался критическим. К сожалению, перечисленные достоинства бесклеточных систем одновременно являются их слабым местом, поскольку после разрушения клеток безвозвратно исчезают те многочисленные взаимодействия между их компонентами, благодаря которым можно без труда отличить живую клетку от бесклеточного экстракта.

Бесклеточные белоксинтезирующие системы представляют собой одни из самых сложных и многокомпонентных систем *in vitro*, используемых в биохимии. Это связано с необходимостью воспроизведения в пробирке всех этапов биосинтеза белка, включая транскрипцию, аминокислотилирование тРНК и трансляцию мРНК рибосомами. Тем не менее, уже более 30 лет удается успешно осуществлять процесс биосинтеза белка *in vitro* и использовать такие системы по трем основным направлениям: для анализа кодирующего потенциала нуклеиновых кислот, исследования механизмов биосинтеза белка и препаративной наработки некоторых рекомбинантных белков и пептидов.

По природе компонентов, которые определяют способность бесклеточных систем осуществлять трансляцию определенных мРНК, принято различать прокариотические и эукариотические системы. Естественно, что наиболее эффективная трансляция мРНК происходит в гомологичных бесклеточных системах. Однако под контролем гомологичных регуляторных элементов в бесклеточных системах могут быть успешно транслированы и чужеродные мРНК – транскрипты рекомбинантных генов.

В некоторых бесклеточных системах транслируют предварительно очищенную мРНК или используют эндогенную мРНК, присутствующую в полисомах. В других белоксинтезирующих системах – системах сопряженной транскрипции и трансляции, одновременно происходят синтез мРНК и ее трансляция рибосомами.

Из-за высокой стоимости очищенных компонентов в подавляющем большинстве случаев бесклеточные системы биосинтеза белка используют в аналитическом варианте, когда объем пробы не превышает 100 мкл. Однако в последнее время разработаны проточные белоксинтезирующие системы, в которых процесс трансляции удается поддерживать длительное время, непрерывно подводя извне расходуемые вещества, с одновременным удалением синтезируемых белков и продуктов деградации компонентов

системы. Основные принципы получения и функционирования всех разновидностей белоксинтезирующих систем будут рассмотрены ниже.

7.5.1. Прокариотические системы

Среди прокариотических бесклеточных белоксинтезирующих систем наибольшее распространение получили системы на основе экстрактов клеток *E. coli*, хотя основные принципы, используемые для их получения, могут быть применены и для других бактериальных клеток. Прокариотическая, как и любая другая бесклеточная система биосинтеза белка, должна содержать несколько обязательных компонентов: 1) рибосомы и белковые факторы трансляции; 2) матричный полирибонуклеотид (мРНК или синтетические полинуклеотиды); 3) аминоацилированные тРНК; 4) GTP в качестве источника энергии; 5) одновалентные (K^+ или NH_4^+) и двухвалентные (Mg^{2+} или Ca^{2+}) катионы, а также буферные вещества для поддержания pH системы в физиологических пределах.

Одним из наиболее важных условий функционирования бесклеточных систем биосинтеза белка является нативное состояние рибосом и факторов трансляции. В качестве источников этих компонентов в бактериальных бесклеточных системах чаще всего используют экстракты соответствующих клеток. Бактериальные клетки выращивают на богатой питательной среде, суспендируют в буфере и разрушают тем или иным способом. Лизат центрифугируют при 30 000 g. При этом в супернатанте (так называемом S30-экстракте) остаются рибосомы и остальные белковые факторы трансляции, необходимые для функционирования системы. Кроме того, в нем содержатся бактериальные ДНК и мРНК, которые, если от них не освободиться, приводят к неспецифическому синтезу белка в бесклеточной системе.

В том случае, если разрушение бактериальных клеток проводят жесткими методами с применением ультразвука, растирания со стеклянными бусами и т.п., освобождение от эндогенных нуклеиновых кислот становится сложной задачей и требует фракционирования экстрактов с помощью ионообменной хроматографии. При такой хроматографической очистке фракции S100 из нее удаляются все аминоацилированные и свободные молекулы тРНК. Поэтому при реконструировании бесклеточной белоксинтезирующей системы к объединенным фракциям рибосом и S100-

экстракта добавляют препарат очищенной суммарной тРНК.

Экстракты бактериальных клеток S30 и S100, кроме всех необходимых факторов трансляции, содержат и основные компоненты системы транскрипции, включая РНК-полимеразу, фактор терминации транскрипции ρ , CRP-белок и т.п. Поэтому для создания ДНК-зависимой системы сопряженной транскрипции и трансляции, в которой происходит полная экспрессия генов, находящихся под контролем бактериальных регуляторных элементов, как правило, не требуется введения в нее дополнительных белков. Достаточно внесения в систему экзогенной ДНК-матрицы, транскрибируемой бактериальной РНК-полимеразой, а также четырех рибонуклеозидтрифосфатов, чтобы в ней начала активно синтезироваться мРНК и одновременно транслироваться рибосомами. В итоге в таких бесклеточных системах в ряде случаев удается синтезировать высокомолекулярные белки, обладающие ферментативной активностью.

Экстракты бактериальных клеток содержат многочисленные нуклеазы и протеолитические ферменты, которые понижают эффективность синтеза белков *in vitro*, разрушая мРНК и образуемые полипептиды. Для преодоления этих затруднений в бесклеточных системах часто используют экстракты мутантных бактериальных клеток, дефектных по РНКазам и полинуклеотидфосфорилазе, а в сами бесклеточные системы при необходимости вводят ингибитор РНКазы из плаценты человека или ингибиторы протеиназ: лейпептин, пепстатин, химостатин и т.п.

Не существует больших ограничений на первичную структуру мРНК, транслируемых в бактериальных бесклеточных системах. Единственным необходимым условием их эффективной трансляции является отсутствие совершенной вторичной структуры или каких-либо особо прочных кооперативных спиральных участков. Вторичную структуру транслируемых мРНК обычно разрушают кратковременным прогреванием с последующим быстрым охлаждением непосредственно перед внесением ее в пробы.

GTP относится к обязательным компонентам бесклеточной белоксинтезирующей системы. Он не может быть заменен на любой другой из известных рибонуклеозидтрифосфатов и необходим для мРНК-зависимого связывания аминоацил-тРНК с рибосомами и транслокации. Поскольку при трансляции происходит непрерывное расходование GTP, а образующийся при

этом GDP является ингибитором трансляции, в процессе функционирования бесклеточной системы осуществляют постоянную регенерацию GTP. Для этого применяют ATP в качестве донора фосфатных групп, которая, в свою очередь, регенерируется путем переноса фосфатной группы вводимого в бесклеточную систему фосфоэнолпирувата с помощью пируватфосфокиназы. Энергия макроэргических связей ATP используется также при аминокислотилировании тРНК аминокислот-тРНК-синтетазами. В том случае, если создается бесклеточная система сопряженной транскрипции и трансляции, для осуществления синтеза РНК в нее вводятся еще два недостающих рибонуклеозидтрифосфата: UTP и CTP.

Для функционирования бесклеточных белоксинтезирующих систем необходимо обеспечивать в них определенные ионные условия. Наиболее существенным фактором в этом случае является концентрация ионов Mg^{2+} . Достаточно изменения оптимальной концентрации Mg^{2+} в системе на 1–2 мМ, чтобы в ней перестали синтезироваться полипептиды, обладающие ферментативной активностью. При этом влияние ионов Mg^{2+} на суммарное включение аминокислот в синтезируемые полипептидные цепи проявляется в меньшей степени, что, по-видимому, объясняется нарушением точности включения аминокислот в белки при неоптимальных концентрациях ионов Mg^{2+} . В системах *in vitro* ионы Mg^{2+} могут быть заменены на ионы Ca^{2+} и даже Mn^{2+} , а также частично замещены полиаминами: спермидином или спермином, которые благоприятно влияют на трансляцию и образование нативных белков.

Биосинтез белка в бесклеточных системах происходит также в присутствии ионов K^+ или NH_4^+ , оптимальные концентрации которых составляют ~100 мМ. Ионы Na^+ ингибируют трансляцию, а ацетат-анионы в используемых солях предпочтительнее анионов Cl^- . Ионы Mg^{2+} , K^+ и NH_4^+ необходимы для ассоциации субчастиц рибосом и поддержания их в компактной форме, а также обеспечения других их функций.

Несмотря на то что в бесклеточной трансляции чаще всего находят применение системы с использованием S30- и S100-экстрактов с добавленными рибосомами, современный уровень знаний молекулярных механизмов трансляции позволяет получать бесклеточные белоксинтезирующие системы из полностью очищенных компонентов. Однако создание таких систем является трудоемким процессом, поэтому их применяют

только в аналитическом варианте для решения специальных задач.

7.5.2. Эукариотические системы

Несмотря на относительную простоту получения бактериальных белоксинтезирующих систем, их использование ограничивается трансляцией бактериальных и фаговых мРНК или рекомбинантных последовательностей нуклеотидов, находящихся под контролем генетических регуляторных элементов бактерий или бактериофагов. В этой связи возникла необходимость в разработке бесклеточных систем биосинтеза белка, пригодных для непосредственной трансляции мРНК высших организмов.

В настоящее время наиболее часто применяемыми системами такого рода являются белоксинтезирующие системы из ретикулоцитов кроликов, зародышей пшеницы и культивируемых соматических клеток различного происхождения. При этом системы всех трех типов могут быть использованы почти с одинаковым успехом для трансляции разнообразных мРНК и, как правило, такие системы не обнаруживают видоспецифичности. Основные принципы конструирования прокариотических белоксинтезирующих систем применяют и для получения систем биосинтеза белков высших организмов. Специфика последних связана, прежде всего, с биологическими особенностями объектов, которые служат источником белковых компонентов систем, а также с различиями механизмов биосинтеза белка у прокариот и эукариот.

Ретикулоциты – безъядерные предшественники эритроцитов человека и животных, осуществляют активный синтез гемоглобина и в связи с этим обладают всеми необходимыми компонентами белкового синтеза. В то же время отсутствие ядер в таких клетках дает возможность получать бесклеточные экстракты, минимально загрязненные геномной ДНК. Эти особенности ретикулоцитов животных делают их излюбленным источником бесклеточных экстрактов, способных осуществлять эффективную трансляцию разнообразных экзогенных мРНК.

Все основные компоненты, необходимые для функционирования бактериальных бесклеточных систем биосинтеза белка, должны присутствовать и в бесклеточной белоксинтезирующей системе из ретикулоцитов кроликов. Однако необходимо отметить две существенные особенности такой системы. Во-первых, для регенерации АТФ в качестве доноров фосфатных групп

используют креатинфосфат – универсальный аккумулятор энергии в макроэргических связях у большинства позвоночных, а сам перенос групп осуществляется вводимой в систему креатинфосфокиназой. Во-вторых, для предотвращения ингибирования трансляции в процессе синтеза белка в систему вводят гемин (механизм его стимулирующего действия обсуждался в первой части книги).

Функционирование системы из зародышей пшеницы и ее состав существенно не отличаются от таковых ретикулоцитарной системы биосинтеза белка. Различия заключаются главным образом в методах получения бесклеточных экстрактов. Зародыши обычно выделяют из сухих зерен озимой пшеницы. Преимуществами этих бесклеточных белоксинтезирующих систем перед другими являются возможность длительного хранения и доступность исходного биологического материала (зерен сортовой пшеницы), что обеспечивает воспроизводимость получаемых результатов и не требует проведения экспериментов с лабораторными животными.

Несмотря на то что разные эукариотические бесклеточные белоксинтезирующие системы активно транслируют гетерологичные мРНК, гомологичные матрицы, как правило, транслируются более эффективно. Например, глобиновые мРНК используются в синтезе белка в 7–8 раз лучше лизатами ретикулоцитов, чем экстрактами культивируемых клеток яичников китайских хомячков. Кроме того, в клетках разных типов дифференцированных тканей могут присутствовать специфические белковые ингибиторы трансляции определенных мРНК. Факты такого рода необходимо учитывать при выборе бесклеточной системы для решения конкретных экспериментальных задач.

7.5.3. Проточные системы

Бесклеточные системы биосинтеза белка позволили генной инженерии получать экспрессию изолированных генов, не прибегая к помощи живых клеток. До недавнего времени все обсуждавшиеся выше бесклеточные системы существовали только в аналитическом варианте и были функционально активны на протяжении короткого времени (40–60 мин). Для всех описанных выше систем характерен малый выход синтезируемого белка (не более 1 мкг на 1 мл экстракта), и на 1 молекулу мРНК в них синтезируется лишь одна–две молекулы полипептида.

Значительный прогресс в этой области исследований был достигнут в лаборатории А.С. Спирина в 1988 г. Здесь с помощью простых усовершенствований удалось получить эффективную бесклеточную белоксинтезирующую систему. В их модификации бесклеточные экстракты бактериальных или эукариотических клеток помещают в ячейку, закрытую с двух сторон полупроницаемыми мембранами. Размер пор позволяет проходить через мембраны вместе с током жидкости низкомолекулярным химическим веществам и небольшим белкам. Содержимое ячейки, в которой имеются все компоненты, необходимые для бесклеточной трансляции, инкубируют при обычной температуре. При этом с одной стороны в такую ячейку-реактор со скоростью ~ 1 мл/ч непрерывно поступают ингредиенты, расходуемые в процессе биосинтеза белка (аминокислоты, АТФ, GTP), а с другой – из нее выходят синтезированные белковые продукты (если их молекулярная масса и отсутствие способности к агрегации позволяют пройти через поры мембраны).

Такое простое усовершенствование бесклеточной системы сильно изменило характер ее функционирования. Биосинтез белка в ней может продолжаться непрерывно в течение нескольких десятков часов, причем на одну молекулу транскрибируемой мРНК синтезируются сотни копий полипептидных цепей белков, суммарный выход которых может достигать 200 мкг и более на 1 мл бесклеточного экстракта.

С разработкой проточной бесклеточной белоксинтезирующей системы стало возможным проводить биосинтез рекомбинантных белков в препаративных количествах *in vitro*. Такая система позволяет синтезировать в больших количествах полипептиды, подверженные быстрой внутриклеточной деградации протеиназами или образующие тельца включения в живых клетках. Проточная система может быть полезна также для получения белков, обладающих цитотоксической активностью, или других рекомбинантных полипептидов, для которых нежелательны артефактные посттрансляционные модификации, происходящие *in vivo*. Недавно в таких системах удалось получить препаративные количества функционально активного интерлейкина 6 человека.

7.6. Другие современные методы исследования генов

Основным методическим достижением генной инженерии в исследовании генов является разработка способов выделения индивидуальных генов и экспрессии их в новом генетическом окружении в гомологичных и гетерологичных генетических системах. Именно эти аспекты генной инженерии были подробно рассмотрены в предыдущих разделах. Однако для изучения тонкой структуры уже клонированных генов, их картирования на хромосомах и исследования механизмов регуляции экспрессии были разработаны другие не менее эффективные методы.

7.6.1. Рестрикционное картирование генов

Полную, но, к сожалению, пока трудно интерпретируемую информацию о строении гена может дать только определение его первичной структуры, т.е. последовательности составляющих ген нуклеотидов. На практике при исследовании протяженных (до 40 т.п.о.) клонированных последовательностей нуклеотидов, включающих исследуемые гены, прежде всего строят их *рестрикционные карты*. Рестрикционные карты представляют собой схемы, изображающие взаимное расположение сайтов рестрикции для разных рестриктаз и расстояния между ними. Поскольку каждый сайт рестрикции является не чем иным, как строго определенной последовательностью нуклеотидов ДНК, рестрикционные карты наглядно заключают в себе информацию об особенностях первичной структуры картируемых участков генома.

Для построения рестрикционной карты используют гибридизацию по методу Е. Саузерна. Клонированный фрагмент ДНК отдельно или в составе вектора получают в препаративном количестве, затем его обрабатывают соответствующими рестриктазами и продукты рестрикции разделяют электрофорезом в агарозном геле. Количество образовавшихся рестрикционных фрагментов ДНК, обнаруживаемых после окрашивания бромистым этидием в виде флуоресцирующих полос в ультрафиолетовом свете, соответствует количеству сайтов рестрикции в том случае, если различия в размерах образовавшихся фрагментов ДНК достаточны для их разделения при электрофорезе.

Размеры рестрикционных фрагментов оценивают путем сравнения их электрофоретической подвижности с таковой фрагментов ДНК известных размеров. Получив информацию о количестве сайтов рестрикции в гене, далее определяют их взаимное расположение. Для этого в качестве зондов выбирают короткие фрагменты ДНК и после введения в них радиоактивной метки их гибридизуют с рестрикционными фрагментами ДНК, которые после электрофоретического разделения в агарозном геле были перенесены на нитроцеллюлозные или нейлоновые фильтры. По завершении гибридизации положение фрагментов ДНК, связавших метку, на фильтрах обнаруживают с помощью автордиографии. Получение такой информации о принадлежности конкретных фрагментов ДНК, образовавшихся под действием различных рестриктаз, к 5'- или 3'-концевым частям исследуемой последовательности нуклеотидов обычно бывает достаточным для определения взаимного расположения различных сайтов рестрикции на рестрикционных картах.

7.6.2. "Прогулки и прыжки по хромосомам"

Многие гены в геноме эукариот располагаются в виде кластеров, члены которых функционально или эволюционно связаны друг с другом. В связи с этим в молекулярной генетике часто возникает задача клонирования генов, расположенных по соседству с уже выделенными генами. Для ее решения был разработан метод *"прогулки по хромосоме"*. Из клонотеки генов, полученной на основе фаговых или космидных векторов, выделяют ген и короткие 5'- и 3'-концевые фрагменты его последовательности используют в качестве зондов для поиска фрагментов ДНК, перекрывающихся с этим геном. Основным требованием, предъявляемым к зонду, является принадлежность его к уникальной последовательности нуклеотидов, т.е. встречающейся в геноме только один раз. С использованием такого простого подхода были исследованы многие кластеры эукариотических генов, в частности β -глобинового локуса человека, структура которого представлена на рис. II.14.

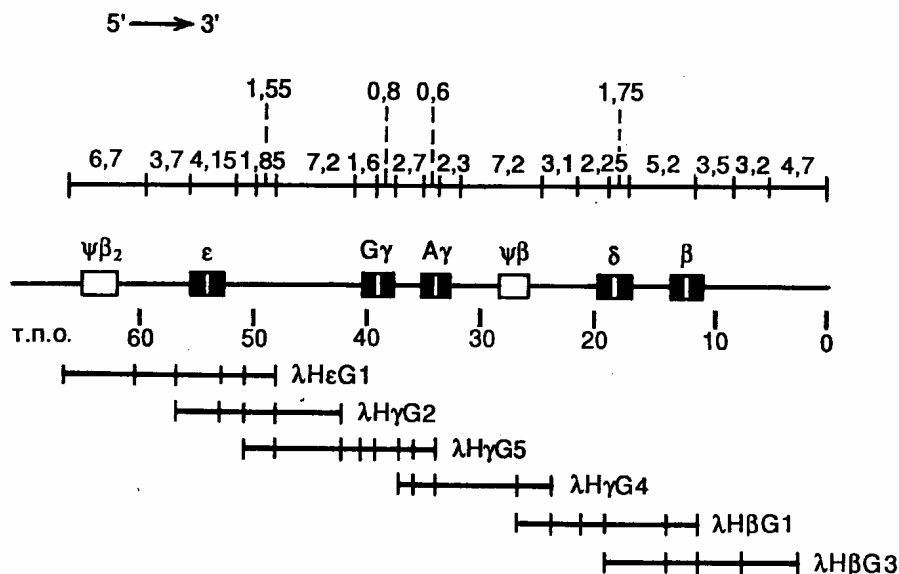


Рис. II.14. Генетическая карта локуса β -глобинового гена человека, полученная методом "прогулки по хромосоме"

Перекрывающиеся фрагменты геномной ДНК человека получены путем ее частичного расщепления рестриктазой *EcoRI*, сайты рестрикции для которой отмечены поперечными черточками. Указаны расстояния между *EcoRI*-сайтами рестрикции в т.п.о., гены β -глобинового локуса, а также названия рекомбинантных фагов λ .

С помощью "прогулки по хромосоме" удастся детально картировать участки ДНК длиной до 250 т.п.о. Недавно разработанный метод "прыжков по хромосоме" позволяет стыковать за один прием фрагменты ДНК общей длиной от 100 до 500–1000 т.п.о. В этом случае на первом этапе препарат высокомолекулярной ДНК расщепляют соответствующей рестриктазой и с помощью электрофореза в импульсном электрическом поле получают фракцию фрагментов ДНК одинакового размера (в данном примере ~100 т.п.о.) (см. рис. II.15). Затем фрагменты ДНК лигируют с маркерным геном, например геном устойчивости к антибиотикам, что приводит к образованию кольцевых молекул. Кольцевые молекулы ДНК расщепляют другой рестриктазой, для которой отсутствуют сайты рестрикции в маркерном гене, что сопровождается образованием коротких фрагментов ДНК, которые далее можно клонировать обычными способами. Полученную клонотеку фрагментов ДНК исследуют с использованием гибридизации с зондом, комплементарным точке начала "прыжка по хромосоме". Отобранные в результате гибридизации клоны

содержат маркерный ген, фланкированный изучаемыми последовательностями нуклеотидов, удаленными друг от друга на 100 т.п.о. Таким образом, получают информацию о последовательностях нуклеотидов, удаленных от точки начала "прыжка по хромосоме" на 100 т.п.о., и эти последовательности далее используют в качестве зонда для выделения и исследования фланкирующих ее фрагментов ДНК.

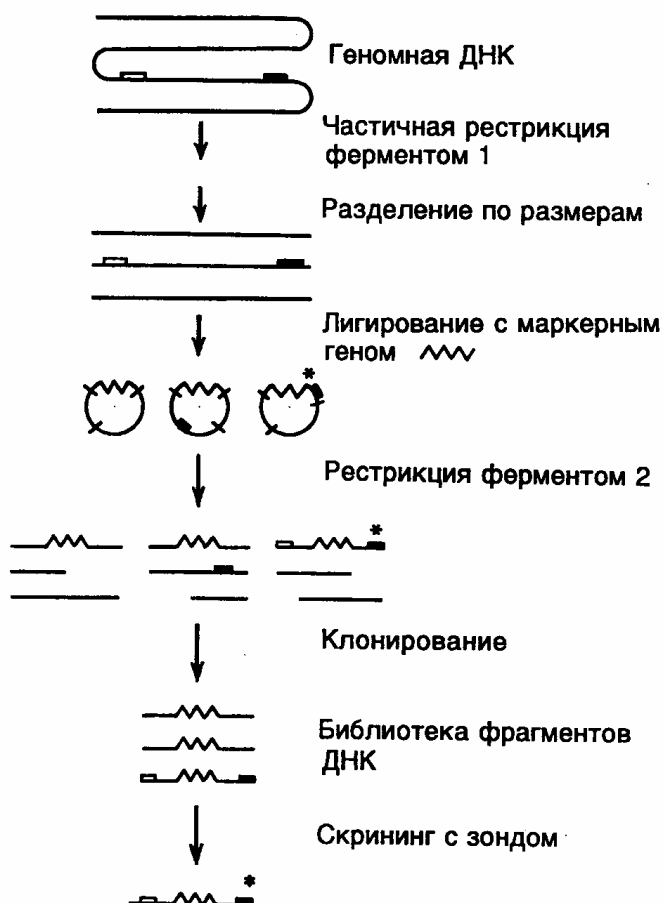


Рис. II.15. Схема метода "прыжков по хромосоме"

7.6.3. S1-картирование РНК и ДНК

Нуклеаза S1, специфически гидролизующая одноцепочечные ДНК и РНК, успешно используется для исследования колинеарности ДНК и кодируемой ей РНК, точного картирования мест инициации и терминации транскрипции на ДНК-матрицах, а также для оценки гомологии между двумя молекулами ДНК

или РНК. Для обнаружения интронов в изучаемых генах клонированный в составе геномной ДНК ген гибридизуют со зрелой РНК, кодируемой этим геном. При наличии в гене интронов их последовательности в ДНК-РНК-гибридах освобождаются в виде одноцепочечных участков и могут быть специфически гидролизованы S1-нуклеазой. При электрофоретическом разделении продуктов гидролиза в денатурирующих условиях при наличии интронов наблюдают появление фрагментов ДНК, причем на основании размеров фрагментов можно локализовать положение интронов в гене. Такой же подход используется и для анализа гомологии между двумя последовательностями нуклеиновых кислот или обнаружения мутаций.

Использование S1-нуклеазы позволяет с точностью до одного–двух нуклеотидов картировать положение точек инициации и терминации транскрипции внутри гена. И в этом случае короткие меченые фрагменты 5'-или 3'-концевых частей гена, заключающие в себе указанные последовательности ДНК, гибридизуют с исследуемыми РНК и образовавшиеся ДНК-РНК-гибриды инкубируют с S1-нуклеазой. Электрофоретическая подвижность полученного фрагмента ДНК, защищенного от действия S1-нуклеазы концевой частью анализируемой РНК, сравнивается с таковой фрагментов того же участка ДНК, образующихся при его секвенировании по методу Максама-Гилберта. В результате однозначно определяют положение нуклеотидов во фрагментах исследуемых ДНК, которые защищены в ДНК-РНК-гибридах от действия S1-нуклеазы.

7.6.4. Футпринтинг

Принцип защиты последовательности нуклеотидов рестрикционных фрагментов ДНК белками от действия агентов, расщепляющих ДНК, лежит в основе футпринтинга – метода, позволяющего определять места специфических контактов белков с ДНК. Этот метод оказался особенно полезным для точной локализации последовательностей нуклеотидов генов, взаимодействующих с регуляторными белками, активаторами и репрессорами, а также с самими РНК-полимеразами, и своему появлению целиком обязан генной инженерии. Рестрикционные фрагменты ДНК, последовательности нуклеотидов которых известны, метят по одному из концов ^{32}P и инкубируют с исследуемыми белками. Образовавшиеся комплексы белок–ДНК подвергают

действию агентов, гидролизующих ДНК, например ДНКазы I, в условиях *неполного расщепления ДНК*, и полученные продукты гидролиза разделяют с помощью электрофореза в полиакриламидном геле с последующей автордиографией. В отсутствие белка на автордиограммах наблюдают появление полного набора фрагментов ДНК в том виде, как это имеет место при обычном секвенировании ДНК. Участки ДНК, защищенные белком от действия ДНКазы, идентифицируются по исчезновению полос, соответствующих продуктам статистического расщепления ДНК, концы которых попадают в область связывания с белком. При этом на электрофоретических дорожках появляются характерные пропуски (или "следы" – footprints), что и дало название всему методу.

7.7. Стратегия выделения нового гена

После обсуждения основных экспериментальных приемов, используемых в современной генной инженерии, становится ясно, каким образом можно решить одну из основных методических задач молекулярной генетики, а именно: выделить требуемый ген и заставить его работать в новых для него генетических условиях. Такая задача наиболее просто решается в случае бактериальных и вирусных генов, поскольку геном этих микроорганизмов невелик. Кроме того, бактериальные гены, как правило, не содержат интронов и даже гетерологичные бактериальные гены хорошо экспрессируются в клетках *E. coli*. Но и клонирование эукариотических генов в настоящее время представляется вполне посильной задачей. Приняв решение о клонировании какого-либо эукариотического гена, прежде всего необходимо ответить на вопрос: нужна ли экспрессия рекомбинантного гена в клетках микроорганизмов или же можно ограничиться исследованием его структуры? Очевидно, что в первом случае нужно думать о создании клонотeki безинтронных кДНК, а во втором – клонотeki геномной ДНК исследуемого объекта, что в ряде случаев технически менее сложно. Задача становится более трудной, если отсутствуют сведения о первичной структуре клонируемого гена. Тогда единственным источником получения частичной информации о последовательности его нуклеотидов может быть кодируемый этим геном белок.

Определив последовательность нескольких N-концевых аминокислот белка, можно в соответствии с генетическим кодом синтезировать ряд

олигонуклеотидных зондов, которые далее используют для скрининга клонотеки генов. В этом случае удобнее использовать экспрессирующую клонотеку кДНК, так как для получения необходимой информации нужно будет исследовать меньшее количество клонов. Действительно, в клонотеках кДНК отсутствует большинство некодирующих последовательностей нуклеотидов, суммарная длина которых может значительно (на два–три порядка) превышать таковую значимых последовательностей. Однако при использовании таких клонотек особенно остро встает вопрос об их репрезентативности, поскольку внутриклеточное содержание индивидуальных мРНК сильно различается в клетках разных тканей. После выделения рекомбинантной ДНК, гибридизующейся с упомянутыми выше олигонуклеотидными зондами, можно идентифицировать кодируемый кДНК белок после ее экспрессии с использованием специфических антител. Альтернативно: кодирующий потенциал клонированной кДНК определяют после ее гибридизации с суммарной мРНК, выделяя фракцию мРНК, задерживаемой иммобилизованной кДНК на носителе, с последующей трансляцией мРНК в бесклеточной белоксинтезирующей системе. Кроме того, если клонированная кДНК кодирует мРНК исследуемого белка, то они и в растворе образуют специфический ДНК-РНК-гибрид, и мРНК перестает участвовать в трансляции в бесклеточной системе (метод прерванной трансляции). Наличие или отсутствие белкового продукта бесклеточной трансляции легко обнаружить с помощью специфических антител. После проведения таких исследований можно точно идентифицировать клонированную кДНК и использовать ее в качестве зонда для выделения целого гена из клонотеки геномной ДНК. Одновременно возможна экспрессия клонированной кДНК в клетках микроорганизмов или в гомологичных эукариотических клетках.

По-прежнему уникальной задачей, которая по плечу лишь большим международным коллективам, остается клонирование неизвестных генов животных и растений, функционирование которых можно заметить по сложным фенотипическим признакам, например симптомам системного наследственного заболевания. При выделении таких генов из генома человека работа начинается с проведения скрупулезного анализа сцепления исследуемого фенотипического признака и какого-либо полиморфного молекулярного маркера, например редкого аллеля HLA или минисателлитного локуса всех

членов семьи больного. После обнаружения сцепленных маркеров задача сводится к клонированию крупных фрагментов ДНК, включающих эти маркеры, их секвенированию, выявлению открытых рамок считывания, в которых пытаются обнаружить мутации, отсутствующие у здоровых индивидуумов. В случае удачи дальнейшая работа может проводиться по одной из вышеприведенных схем и должна закончиться идентификацией гена и его белкового продукта.

Примерно такой путь исследований в генной инженерии привел за последние 20 лет к клонированию множества генов, в том числе и неизвестных ранее, определению их структуры и особенностей функционирования. Аналогичные подходы легли в основу новых направлений молекулярной генетики. О некоторых из них речь пойдет ниже в других главах книги.

ГЛАВА 8. НАПРАВЛЕННЫЙ МУТАГЕНЕЗ И БЕЛКОВАЯ ИНЖЕНЕРИЯ

Многие мутации сопровождаются изменениями фенотипа мутантного организма на морфологическом или биохимическом уровне, хорошо различимыми визуально или с помощью биохимических методов. Такие мутационные изменения называют *генетическими маркерами* мутантного организма, так как они позволяют отличать мутанты от нормальных особей и проследивать передачу мутантных генов в ряду поколений. Именно эта возможность следить за судьбой мутантных генов в потомстве, получаемом от скрещиваний родительских мутантных организмов, позволила в свое время создать мощное научное направление, называемое *формальной генетикой*, основанное на *генетическом анализе* наследования мутантных признаков.

До развития геной инженерии гены исследовали преимущественно методами генетического анализа с использованием мутантных признаков организмов в качестве генетических маркеров. Это давало возможность определения границ генов и построения их подробных генетических карт. Именно потребности генетического анализа привели к быстрому развитию классических методов получения мутаций: чем больше имеется мутаций в определенном гене, тем более подробную генетическую карту такого гена можно построить. Классические методы получения мутаций в конкретных генах основаны на отборе мутантных организмов в селектируемых условиях (например по устойчивости к антибиотикам) из большого числа организмов дикого типа и мутантных организмов, несущих мутации в других генах. После получения искомого мутантного организма можно локализовать мутацию на генетической карте и биохимическими методами исследовать ее проявление на молекулярном уровне. Например, с использованием очищенного мутантного фермента можно определить молекулярные изменения, приводящие к формированию мутантного фенотипа. Этот традиционный подход, не забытый окончательно и сегодня, дал очень много для развития классической и молекулярной генетики. Однако из-за низкой эффективности и большой трудоемкости метода его было трудно использовать для исследования высших организмов.

8.1. Методы направленного получения мутаций

Развитие генной инженерии революционизировало процесс получения мутаций в конкретных участках генома и анализ последствий этих мутаций на молекулярном уровне. Совокупность методов получения мутаций, основанных на использовании генно-инженерных подходов, называют *направленным*, или *сайт-специфическим мутагенезом*. Протообразом направленного мутагенеза является простая обработка клонированного гена химическими мутагенами *in vitro*, которая приводит к четкой локализации возникающих мутаций в пределах этого гена. В настоящее время разработано много эффективных методов сайт-специфического мутагенеза, позволяющих сознательно производить мутационные замены конкретных нуклеотидов. Направленное получение мутаций в сегментах рекомбинантных генов, введение мутантных генов в организм и исследование влияния полученных мутаций на функционирование гена – этот подход прямо противоположен классическим методам получения мутаций. Совокупность таких подходов получила название обратной генетики. Кроме того, возможность замены конкретных аминокислот в белках с известной первичной структурой, а также объединение в одной полипептидной цепи доменов различных белков и ферментов позволили, по сути дела, конструировать *in vitro* новые белки, не встречающиеся в природе, и привели к созданию в молекулярной генетике нового направления – *белковой инженерии*. В следующих разделах этой главы будут рассмотрены основные методы направленного мутагенеза и белковой инженерии, а также некоторые впечатляющие результаты, полученные с применением этих методов.

8.1.1. Получение делеций и вставок

Делецией называют потерю части нуклеотидов в геноме организма. Такой вид мутаций удобнее всего использовать для локализации (картирования) функционально значимых участков генов и кодируемых этими генами белков. Действительно, последовательное удаление все новых и новых участков ДНК на границах генов с помощью делеций оказалось исключительно плодотворным в обнаружении регуляторных элементов генов, исследовании их структурно-функциональных особенностей, взаимного расположения и влияния друг на друга. Простой и эффективный метод получения делеций любого

размера разработан с использованием экзонуклеазы *Bal31* для удаления нуклеотидов в окрестностях сайтов рестрикции (рис. II.16).

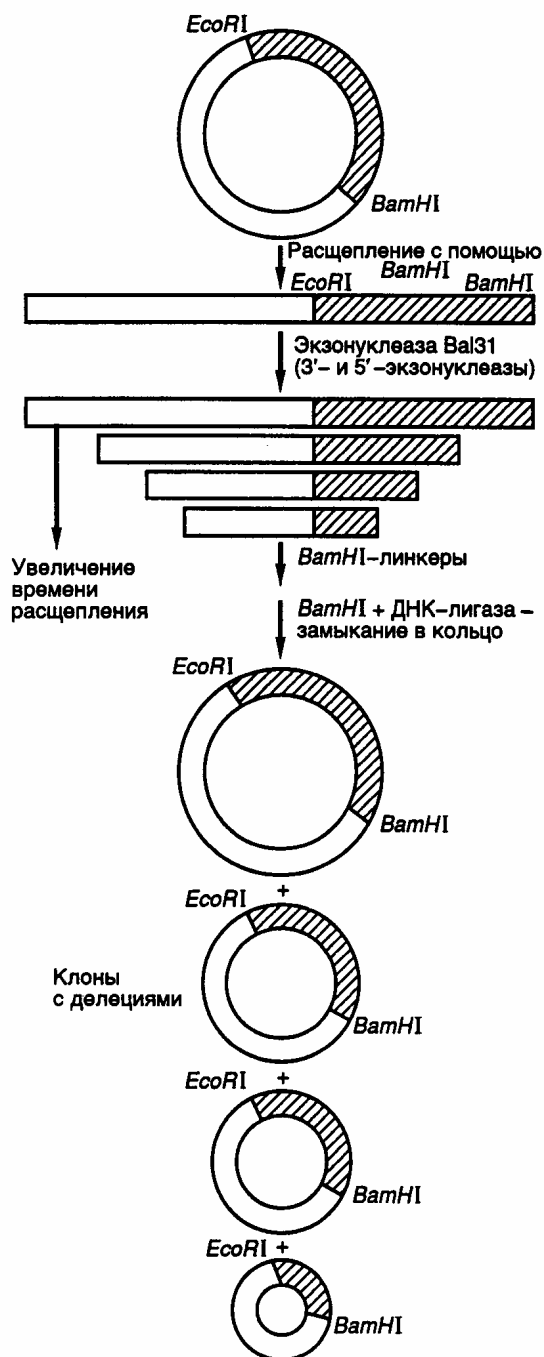


Рис. II.16. Получение делеций с помощью нуклеазы *Bal31*

Ген, клонированный в плазмиде, расщепляют по уникальному сайту рестрикции и образовавшиеся линейные молекулы ДНК инкубируют в присутствии экзонуклеазы Bal31. При этом экзонуклеаза последовательно удаляет нуклеотиды с обоих концов ДНК, причем количество удаляемых нуклеотидов прямо пропорционально времени инкубации ДНК с нуклеазой, а также зависит от температуры инкубации и концентрации фермента. В результате подобного действия образуется набор фрагментов ДНК разной длины, содержащих делеции различных размеров по обе стороны выбранного сайта рестрикции. К "тупым" концам таких молекул ДНК с помощью ДНК-лигазы присоединяют двухцепочечные олигонуклеотидные линкеры, содержащие уникальный (часто исходный) сайт рестрикции, обрабатывают соответствующей рестриктазой, замыкают молекулы в кольцо посредством лигирования и затем вводят их в бактериальные клетки. Точное картирование концов делеций осуществляют секвенированием соответствующих участков ДНК мутантных плазмид. В результате получают набор делеций разного размера, положение которых в исследуемом фрагменте ДНК строго локализовано.

Небольшие делеции в окрестностях сайтов рестрикции можно получать быстрее удалением "липких" концов линеаризованной плазмидной ДНК после рестрикции с последующим замыканием линейной ДНК в кольцо лигированием по образовавшимся "тупым" концам. В этом случае размер делеции соответствует размеру одноцепочечных "липких" концов в сайтах рестрикции. Кроме того, такой метод допускает простую проверку наличия мутаций в требуемом участке ДНК, так как в результате мутации происходит потеря уникального сайта рестрикции.

Для получения вставок коротких или протяженных последовательностей нуклеотидов в исследуемые участки клонируемых генов также разработаны эффективные и надежные методы. В простейшем случае такая задача решается путем расщепления исследуемого гена рестриктазой по уникальному сайту рестрикции и встраивания по этому сайту фрагмента природной ДНК или синтетического двухцепочечного олигонуклеотида, который фланкирован соответствующими "липкими" концами. Лигирование может проводиться и по "тупым" концам. В таком случае последовательности нуклеотидов на концах вставки уже не имеют существенного значения для встраивания этого фрагмента по сайту рестрикции.

Для создания множественных вставок коротких или протяженных последовательностей нуклеотидов в исследуемых участках ДНК в основном используют два подхода. В первом случае с помощью панкреатической ДНКазы в низких концентрациях в присутствии ионов Mn^{2+} вносят случайным образом один двухцепочечный разрыв в каждую векторную плазмиду, содержащую клонированный ген. К концам образовавшихся линейных молекул ДНК присоединяют с помощью ДНК-лигазы синтетические олигонуклеотидные линкеры, содержащие сайт рестрикции, который отсутствует в исследуемой плазмиде. Образовавшиеся линейные молекулы ДНК с линкерами обрабатывают рестриктазой, узнающей сайт рестрикции линкера, что приводит к образованию "липких" концов, и замыкают в кольцо с помощью ДНК-лигазы. В итоге кольцевые молекулы ДНК содержат исследуемый клонированный ген, в котором имеется по одной вставке локализованных случайным образом (в соответствии с расположением исходных двухцепочечных разрывов) олигонуклеотидных линкеров. Во втором случае статистические разрывы в двухцепочечной ДНК получают путем частичного (неполного) гидролиза мелкощепящими рестриктазами, которые узнают сайт рестрикции длиной в 4 нуклеотида. Метод получения вставок с использованием синтетических олигонуклеотидных линкеров получил название *сканирования линкером*.

8.1.2. Химический мутагенез

Делеции и вставки, создаваемые в структурных частях генов, как правило, их инактивируют, особенно в тех случаях, когда такие мутации приводят к сдвигу открытых рамок считывания. Поэтому делеции и вставки *in vitro* используют главным образом для поиска и изучения регуляторных элементов генов, влияющих на эффективность их экспрессии. Большое значение для исследования функционирования белков имеют методы мутагенеза *in vitro*, направленные на получение точковых мутаций, следствием которых являются одиночные замены аминокислот в полипептидных цепях.

Распространенным методом введения большого числа точковых мутаций разной локализации в исследуемые части генов *in vitro* является химический мутагенез одноцепочечных участков рекомбинантных ДНК. Принцип подобных методов заключается в том, что некоторые химические мутагены, такие как бисульфит натрия, гидроксилламин или метоксиламин, действуют только на

одноцепочечные участки ДНК. Следовательно, получив молекулы ДНК, содержащие одноцепочечные бреши в исследуемых участках генов, можно с помощью бисульфита натрия дезаминировать остатки цитозина в этих участках, т.е. превратить их в остатки урацила. После достройки цепи такой мутагенизированной молекулы ДНК с помощью фрагмента Кленова ДНК-полимеразы I *E. coli* происходит замена исходных G-C-пар на T-U. Затем мутагенизированные молекулы ДНК с помощью трансформации вводят в бактериальные клетки, где по завершении первого раунда репликации в молекуле осуществляется замена остатков U на T и полная замена G-C-пары на A-T, т.е. имеет место обычная транзиция.

Одноцепочечные мутагенизированные участки ДНК удобно получать путем гибридизации одноцепочечной ДНК вектора с двухцепочечной ДНК того же вектора, содержащего клонированный ген или его участок, который необходимо мутагенизировать. В этом случае в образующемся гибриде-гетеродуплексе, одна цепь которого принадлежит вектору без вставки, а другая – вектору со вставкой, происходит выпетливание последовательности вставки в виде одноцепочечного участка ДНК. Обсуждаемый подход к получению статистического набора точковых мутаций с использованием химических мутагенов позволяет легко создавать большое число мутантных молекул ДНК, содержащих одну или несколько мутаций в разных сочетаниях. Последующий отбор мутантов на основе новых биохимических или иных параметров мутантных белков (исчезновение, ослабление или усиление ферментативной активности, появление новой активности или новых иммунологических свойств и т.п.) позволяет идентифицировать остатки аминокислот в исследуемых белках, отвечающие за эти изменения.

Несмотря на удобство введения такого рода мутаций в ДНК *in vitro*, химический мутагенез накладывает ограничения на спектр возникающих мутаций, так как лишь определенные остатки нуклеотидов ДНК претерпевают обязательные изменения. Поэтому многие мутации никогда не могут быть получены с помощью химических мутагенов. Проблему можно частично решить, используя для репарации одноцепочечных брешей ДНК аналоги нуклеотидов, например N-гидроксицитозинтрифосфат, который в составе ДНК одинаково хорошо спаривается с A и G, или создавая такие условия, при которых репарирующая ДНК-полимераза начинает ошибочно включать в

синтезируемую цепь ДНК некомплементарные матрице нуклеотиды. Все перечисленные выше методы локального мутагенеза, осуществляемого *in vitro*, позволяют, в конечном счете, получать набор случайных мутаций, локализованных на определенном исследуемом участке ДНК. Мутагенизированные молекулы ДНК из одной реакционной пробирки представляют собой сложную смесь, в которой каждая молекула несет несколько независимо возникших мутаций. Для введения мутаций в определенный локус исследуемого гена необходимо проводить трудоемкую процедуру отбора, сопряженную с анализом большого числа мутантов. Подлинную революцию в направленном мутагенезе произвела разработка методов с использованием синтетических олигонуклеотидов.

8.1.3. Сайт-специфический мутагенез с использованием олигонуклеотидов

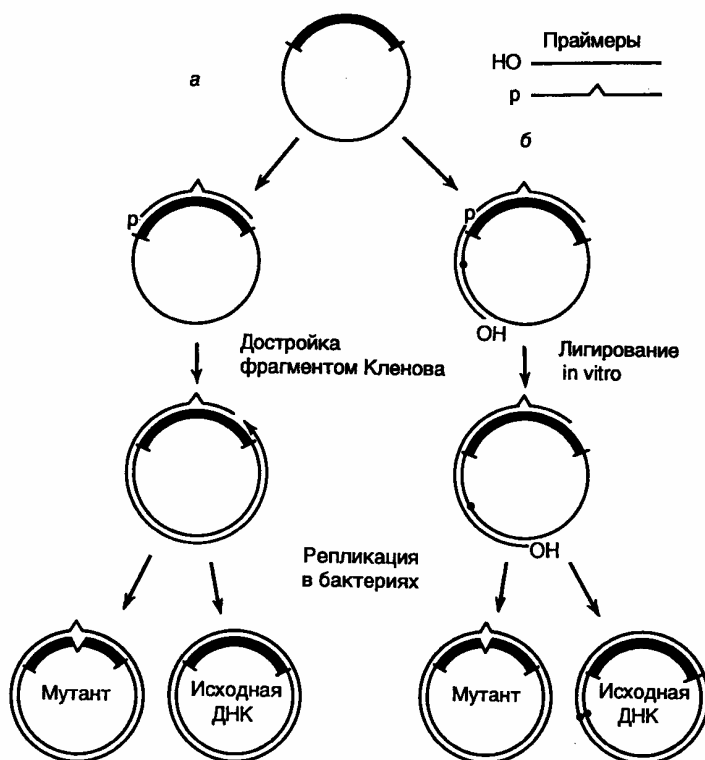


Рис. II.17. Направленный мутагенез с использованием олигонуклеотидов

а – система с одним мутагенизирующим праймером;

б – система с двумя праймерами и селективируемым маркером

В простейшем случае для проведения желаемой замены нуклеотидов синтезируют 15–20-членный олигодезоксирибонуклеотид, комплементарный одной из цепей ДНК исследуемого гена (рис. II.17,а). При этом в данный нуклеотид в процессе синтеза вносят требуемые замены нуклеотидов. Мутагенизируемый ген клонируют в фазмидном векторе, который переводят в одноцепочечную форму. После выделения такой одноцепочечной ДНК с клонированным геном дикого типа с ней гибридизуют синтетический олигонуклеотид и, используя его в качестве затравки, синтезируют с помощью фрагмента Кленова *in vitro* комплементарную цепь ДНК. Образующаяся в итоге двухцепочечная кольцевая молекула ДНК представляет собой гетеродуплекс, дочерняя цепь которого комплементарна исходной родительской цепи на всем протяжении, кроме участка, занимаемого олигонуклеотидом, где мутантный нуклеотид (нуклеотиды) некомплементарен соответствующему нуклеотиду матрицы.

Остающийся после завершения синтеза дочерней цепи ДНК одноцепочечный разрыв ликвидируют с помощью ДНК-лигазы, а образовавшуюся двухцепочечную ковалентно замкнутую ДНК вводят в бактериальные клетки, где она проходит несколько раундов репликации. После первого раунда репликации происходит сегрегация в разные молекулы ДНК мутантного аллеля и аллеля дикого типа исследуемого гена, и далее эти молекулы продолжают реплицироваться независимо друг от друга, что приводит к образованию двух популяций рекомбинантных ДНК, содержащих изучаемый ген в мутантной и исходной формах. Таким образом, теоретически каждая вторая молекула, выделенная из подобных бактериальных клеток, должна содержать в клонированном гене требуемую мутацию, локализованную в строго определенном сайте.

Для того чтобы отличить мутантные ДНК от нормальных, проводят либо прямое секвенирование мутагенизируемого участка гена, либо гибридизацию рекомбинантных молекул с олигонуклеотидом, ранее использованным в качестве затравки, предварительно меченным радиоактивными изотопами (например ^{32}P). Связь этого олигонуклеотида, содержащего замены нуклеотидов, в гибриде будет более прочной с мутантной ДНК, чем с ДНК дикого типа, так как в первом случае олигонуклеотид и ДНК будут полностью комплементарны друг другу. Плавление гибрида и отделение олигонуклеотида

от ДНК произойдут при более низких температурах, если гибрид был образован с ДНК дикого типа, а не с мутантной ДНК. Освободившиеся в результате плавления гибридов олигонуклеотиды отделяют от ДНК промыванием комплекса на фильтрах, что проявляется в отсутствие сигнала при автордиографии, тогда как в случае гибридов олигонуклеотида и мутантной ДНК будет появляться четкий сигнал.

На практике, однако, частота возникновения мутантных ДНК в общей популяции рекомбинантных мутагенизируемых молекул бывает значительно $> 50\%$. Это обусловлено рядом причин, основными из которых *in vitro* являются $3' \rightarrow 5'$ -экзонуклеазная активность фрагмента Кленова и его способность вытеснять цепь ДНК в направлении $5' \rightarrow 3'$. Для преодоления первого затруднения мутантные нуклеотиды в олигонуклеотидном праймере располагают не ближе, чем за четыре нуклеотида от его $3'$ -конца (обычно за семь–десять нуклеотидов). Чтобы избежать последствий второго источника артефактов, для синтеза комплементарной цепи в мутагенизируемой одноцепочечной молекуле ДНК используют систему из двух праймеров (см. рис. II.17,б). В этой системе второй праймер выбирается таким образом, чтобы его $3'$ -конец после гибридизации с мутагенизируемой одноцепочечной молекулой ДНК располагался на расстоянии 1–3 т.п.о. от фосфорилированного $5'$ -конца мутагенизирующего нуклеотида. В этом случае после заполнения бреши между двумя олигонуклеотидами вновь синтезированной ДНК, синтез которой начался с дополнительного праймера, произойдет лигирование одноцепочечного разрыва ДНК-лигазой, которая присутствует в реакционной смеси вместе с фрагментом Кленова, поскольку $5'$ -конец мутагенизирующего нуклеотида фосфорилирован. Однако здесь не образуется кольцевой ковалентно замкнутой молекулы ДНК, так как на $5'$ -конце дополнительного праймера отсутствует фосфатная группа. После завершения реакции образовавшимися молекулами ДНК трансформируют клетки *E. coli* и производят отбор мутантных молекул ДНК в клонах бактерий по гибридизации с мутагенизирующим олигонуклеотидом, меченным радиоактивным изотопом.

Эффективность системы направленного мутагенеза с использованием двух олигонуклеотидов приближается к максимально теоретически возможной – 50% . С помощью этого метода можно получать все виды мутаций, как точковые с заменами одного–трех оснований, так и протяженные делеции длиной в 0,5–

1,0 т.п.о., а также вставки, например для создания новых сайтов рестрикции. Если для получения точковых мутаций достаточно мутагенизирующего праймера длиной в 17–20 нуклеотидов, то в случае получения делеций праймеры должны быть длиннее (30–40 нуклеотидов), поскольку они удерживают последовательности матричной ДНК, фланкирующие выпетливаемый участок. Дальнейшее развитие системы направленного мутагенеза с использованием двух олигонуклеотидов, получившей название *системы сопряженного праймирования*, позволило значительно упростить процедуру отбора мутантных ДНК. В таких усовершенствованных системах в дополнительный олигонуклеотид вводится мутация, инактивирующая или восстанавливающая в дочерней мутантной ДНК какой-либо селектируемый маркер. В одной из таких систем, где в качестве вектора используется ДНК фага M13mp18amIV, селектируемым маркером служит амбер-мутация в гене IV. Матричную ДНК получали, выращивая мутантные фаги M13mp18amIV в бактериальных клетках *E. coli* TG1, содержащих ген-супрессор амбер-мутаций Su2⁺. При этом данный фаг не развивается в клетках несупрессорных штаммов *E. coli* Su2⁻.

Для проведения мутагенеза одноцепочечную ДНК фага M13mp18am1Y гибридизуют с двумя праймерами, один из которых мутагенизирующий, а другой, комплементарный гену IV, содержит точковую мутацию в амбер-кодоне, восстанавливающую последовательность нуклеотидов дикого типа. После достройки комплементарной цепи ДНК фага M13mp18am1Y с использованием вышеупомянутых праймеров двухцепочечные ДНК вводят в бактериальные клетки Su2⁻, которые, кроме того, дефектны по системе репарации. Поскольку бактериальные клетки, дефектные по системе репарации, сами по себе обладают мутаторным фенотипом (т.е. для них характерна повышенная частота возникновения спонтанных мутаций), то для того чтобы в мутируемые ДНК не вносить дополнительные мутации, трансформированные бактериальные клетки высевает на газон клеток *E. coli* Su2⁻ с нормальной системой репарации. При таком подходе фаговые частицы проходят лишь первый цикл развития в клетках с двумя мутациями, а все последующие – в клетках Su2⁻ с одной мутацией. В результате в таких клетках образуются только те фаговые частицы, ДНК которых содержат сайт-специфическую мутацию и

восстановленный ген IV.

Аналогичный подход был использован в системе с геном β -галактозидазы, находящимся в составе векторов серии M13mp. В такой системе фаговый ген кодирует α -пептид β -галактозидазы *E. coli* и комплементирует (восстанавливает) дефектную β -галактозидазу бактериальных клеток-хозяев. Введение амбер-мутации в фаговый ген β -галактозидазы с помощью праймера, используемого в паре с олигонуклеотидом-мутатором, сопровождается потерей дочерними фаговыми частицами способности комплементировать мутантную β -галактозидазу бактериальных клеток. В результате в образующихся фаговых бляшках отсутствует активность β -галактозидазы, которая бы расщепляла искусственный субстрат BCIG с образованием продукта, окрашенного в голубой цвет. Вследствие этого бляшки, образованные фаговыми частицами, содержащими ДНК с сайт-специфическими мутациями, не окрашены и их легко отбирать визуально.

Хорошо себя зарекомендовали и системы циклического отбора мутантных ДНК, содержащие сайт-специфические мутации, в которых используются свойства систем рестрикции–модификации EcoK и EcoB *E. coli*. Используемые для трансформации штаммы *E. coli* могут содержать системы рестрикции–модификации EcoK- или EcoB-типа. В том случае, если родительскую цепь ДНК для мутагенеза выделяют из клеток, дефектных по метилазам системы модификации, то после достройки комплементарной цепи с мутантного олигонуклеотида такая неметилованная ДНК будет расщепляться системой рестрикции после введения в соответствующий бактериальный штамм. Если одновременно с олигонуклеотидом-мутатором использовать второй олигонуклеотид, изменяющий сайт рестрикции, то лишь ДНК с мутировавшим сайтом рестрикции в результате ошибочного функционирования системы репарации сохранится в бактериальных клетках. Такие ДНК с высокой вероятностью будут содержать и требуемую мутацию.

Недавно была описана система, в которой в родительскую цепь ДНК фазмиды вводили вместо остатков тимидина остатки уридина. В этом случае после синтеза дочерней комплементарной ДНК *in vitro* с использованием мутантного олигонуклеотида в качестве затравки родительская цепь ДНК инактивируется в трансформированных бактериальных клетках с помощью

эндогенной урацил-N-гликозилазы, удаляющей из ДНК остатки урацила, что сопровождается частичным гидролизом этой цепи с последующей репарацией брешей по мутантной цепи в качестве матрицы и закреплением введенной мутации.

Наконец, следует упомянуть систему олигонуклеотид-зависимого сайт-специфического мутагенеза, в которой для отбора мутантных цепей ДНК используется попеременно то один, то другой селектируемый маркер, в качестве которых выступают гены устойчивости к антибиотикам. Это позволяет вводить с высокой эффективностью в один и тот же ген несколько сайт-специфических мутаций в последовательных циклах направленного мутагенеза. В таких системах в первом цикле вместе с мутагенизирующим праймером в качестве селектируемого используют праймер, вносящий мутацию, которая придает бактериальным клеткам устойчивость к одному антибиотику, а во втором цикле – к другому (с восстановлением чувствительности к первому антибиотику). Далее такие циклы могут быть повторены необходимое число раз, которое определяется количеством требуемых мутаций.

Разработка систем сопряженного праймирования позволила значительно увеличить эффективность направленного мутагенеза и довести выход ДНК, содержащих сайт-специфические мутации, до 70% от общего числа образующихся молекул ДНК.

Рассмотренные подходы к получению сайт-специфических мутаций с помощью олигонуклеотидов позволяют с высокой точностью и эффективностью производить замены отдельных нуклеотидов в строго определенных локусах. Однако чаще всего невозможно предсказать фенотипические последствия мутационных замен отдельных аминокислот в полипептидных цепях белков, и для получения необходимых фенотипических изменений требуется введение множественных мутаций в определенные участки гена с последующим отбором мутантов требуемого фенотипа. Для решения подобных задач был разработан и эффективно используется метод *кассетного мутагенеза*.

При реализации такого подхода из гена, клонированного в составе векторной плазмиды, по двум близко расположенным уникальным сайтам рестрикции вырезается фрагмент ДНК, в который необходимо внести мутации, и на его место встраивается синтетический двухцепочечный олигонуклеотид,

содержащий необходимые замены нуклеотидов (кассету мутаций). В этом случае, если в окрестностях мутагенизируемого локуса гена отсутствуют подходящие природные сайты рестрикции, их вводят с помощью направленного мутагенеза. Разработка автоматических синтезаторов ДНК сделала синтез олигодезоксирибонуклеотидов простой и даже рутинной процедурой. Более того, использование на определенных этапах синтеза вместо одного нуклеотида смеси из двух, трех или даже всех четырех дезоксирибонуклеозидтрифосфатов позволяет получать за один прием сложную смесь олигонуклеотидов, которые могут содержать в определенных сайтах наборы кодонов для многих или даже для всех 20 природных аминокислот. Это дает возможность осуществлять одновременный скрининг по искомому мутантному фенотипу большого числа разных мутантных клонов, полученных в одном цикле клонирования. С помощью кассетного мутагенеза можно легко исследовать функциональную роль отдельных сайтов и целых доменов в полипептидных цепях конкретных белков и создавать рекомбинантные белки с новыми, подчас неожиданными свойствами.

Внесение множественных мутаций в разные участки полипептидных цепей исследуемых белков позволяет при наличии соответствующих эффективных методов отбора получать белки с требуемыми свойствами и уже после этого определять, какие именно аминокислоты придают такие свойства белку. Альтернативным методом исследования функциональной значимости отдельных аминокислот в белках является их целенаправленная замена на нейтральную аминокислоту, например аланин. Такая последовательная замена аминокислот в полипептидных цепях изучаемых белков на остатки аланина получила название *сканирование аланином*. Введение аланина в полипептидные цепи не изменяет их общей конформации, как это имеет место, например в случае замен на глицин или пролин, и не сопровождается ярко выраженными электростатическими или стерическими эффектами. Кроме того, аланин часто встречается в полипептидных цепях и с одинаковой частотой представлен как на внутренних, так и на внешних участках полипептидных цепей белковых глобул. С помощью сканирования аланином можно локализовать аминокислоты, образующие активный центр ферментов, исследовать участки полипептидных цепей, существенные для взаимодействия белков с другими макромолекулами и низкомолекулярными лигандами, изучать

структуру эпитопов полипептидных цепей, а также ряд других структурных и функциональных особенностей белков.

Еще одним универсальным методом замены определенных аминокислот в полипептидных цепях белков *in vivo* является использование аминоацилированных различными аминокислотами супрессорных тРНК, которые узнают нонсенс-триплеты в мутантных мРНК в процессе трансляции. В результате в соответствующее место полипептидной цепи взамен аминокислоты, присутствующей в белке дикого типа, встраивается аминокислота, которую несет аминоацилированная супрессорная тРНК. В настоящее время в дополнение к природным супрессорным тРНК *E. coli* синтезированы *in vitro* гены, кодирующие супрессорные РНК новой специфичности. В итоге до 13 аминокислотных замен может быть произведено в синтезирующейся *in vivo* полипептидной цепи в результате супрессии кодона UAG (amber). С использованием такого подхода, в частности, удалось произвести >1600 замен аминокислот Lac-репрессора *E. coli* и локализовать участки полипептидной цепи, существенные для связывания индуктора, прочного связывания оператора и термостабильности белка.

Важным преимуществом метода амбер-супрессии перед другими методами направленного мутагенеза у бактерий и бактериофагов является то, что он не требует синтеза большого числа мутантных генов и их последующего отбора, так как введение одного мутантного гена в составе экспрессирующего вектора в клетки разных супрессорных штаммов *E. coli* позволяет получать разные замены аминокислот с одновременной сверхпродукцией мутантного белка в бактериальных клетках. Это облегчает его последующую очистку и изучение биохимических свойств.

8.1.4. Полимеразная цепная реакция в направленном мутагенезе

Разработка метода полимеразной цепной реакции принципиально изменила ситуацию в исследованиях по направленному мутагенезу. Использование ПЦР для направленного мутагенеза основано на применении в качестве праймеров олигонуклеотидов, не полностью комплементарных матричной ДНК. Повышенные требования к комплементарности накладываются лишь на последний, 3'-концевой, нуклеотид праймера, но даже в случае 3'-концевой некомплементарности праймеры часто продолжают функционировать

в системе ПЦР, хотя и с разной эффективностью. В то же время некоторая способность 17-нуклеотидного праймера инициировать ПЦР проявляется даже при наличии всего восьми нуклеотидов, комплементарных матрице, три из которых расположены на его 3'-конце. Таким образом, простейшим способом введения сайт-специфических мутаций в амплифицируемый фрагмент ДНК является использование праймеров, частично комплементарных матрице, т.е. содержащих необходимые мутантные нуклеотиды. Этот способ применяется для создания в амплифицируемом продукте новых сайтов рестрикции для последующего его клонирования. Такой подход удобен для встраивания амплифицируемого продукта в экспрессирующий вектор в одной открытой рамке считывания (ОРС) с инициирующим ATG-кодоном вектора. Действительно, как уже обсуждалось в главе 7, экспрессирующие векторы часто содержат 5'-концевую часть будущего рекомбинантного гена, включая регуляторные последовательности, промотор, ATG-кодон и иногда дополнительные кодоны нескольких последующих аминокислот. При конструировании таких векторов в последовательности, следующие за ATG-кодоном, вводят сайты рестрикции, по которым и встраивают фрагмент экспрессируемого гена, кодоны которого (для полноценной экспрессии гена) должны находиться в одной ОРС с ATG-кодоном вектора. Положение природных сайтов рестрикции в клонируемом гене редко отвечает требованиям ОРС вектора. Поэтому введение искусственных сайтов рестрикции в клонируемый фрагмент с помощью праймеров, не полностью комплементарных матрице, является хорошим решением проблемы.

Вырожденные праймеры, которые представляют собой сложную смесь олигонуклеотидов, содержащих многие точковые мутации, также используют для сканирования мутациями определенных участков ДНК. В таком классическом варианте постановки ПЦР кроме точковых мутаций можно легко получать делеции и вставки. Дальнейшим усовершенствованием метода направленного мутагенеза с помощью ПЦР явилась разработка подхода к получению гибридных генов с помощью перекрывающихся праймеров (рис. II.18). Этот метод позволяет не только целенаправленно получать точковые мутации, делеции и вставки, но и гибридные молекулы ДНК без применения ДНК-лигазы.

Для получения точковых мутаций используют пару праймеров, в которых

изменен соответствующий нуклеотид или небольшая группа нуклеотидов (мутантный сайт обозначен на рис. II.18,1 прямоугольниками внутри праймеров). При этом внутренние праймеры b и c комплементарны друг другу.



Рис. II.18. Использование ПЦР для получения мутаций

Получение точковых мутаций (1), делеций (2) и вставок (3): a и d – внешние праймеры, b и c – внутренние праймеры. АВ и CD – фрагменты ДНК, образовавшиеся в результате ПЦР с использованием праймеров a и b, а также c и d соответственно. После гибридизации друг с другом объединяемых фрагментов ДНК сформировавшимися комплементарными последовательностями нуклеотидов одноцепочечные участки достраиваются ДНК-полимеразой в процессе ПЦР

В первых ПЦР в отдельных пробирках амплифицируют правую и левую части мутагенизируемого сегмента ДНК с использованием пар праймеров $a + b$ и $c + d$, что приводит к образованию двух фрагментов ДНК, перекрывающихся на участках, которые содержат включенные праймеры (этапы 1 и 2). Фрагменты очищают от праймеров, смешивают друг с другом в эквимольном соотношении и после цикла тепловой денатурации и ренатурации используют в качестве матрицы в другой ПЦР с внешними праймерами a и d (этап 3). На этом этапе в первом цикле ПЦР происходит достройка цепей перекрывающихся фрагментов ДНК (пунктирные линии на рис. II.12,1), и образовавшийся фрагмент двухцепочечной ДНК, содержащий требуемую мутацию, далее служит матрицей для амплификации и по завершении ПЦР присутствует в реакционной смеси в препаративном количестве.

Тот же принцип объединения двух фрагментов ДНК, перекрывающихся за счет праймеров, используется при получении делеций и вставок. Для создания делеции (на рис. II.18,2 обозначена прямоугольниками на цепях ДНК) берут внутренние праймеры b и c , 5'-концы которых комплементарны матрице по одну сторону делетируемой последовательности нуклеотидов, а 3'-концы – последовательности нуклеотидов, фланкирующей другой ее конец. Образующиеся при этом продукты ПЦР АВ и CD перекрываются в точке делеции и далее используются, как и в предыдущем случае. При получении вставок (см. рис. II.18,3) применяют внутренние праймеры b и c , комплементарные друг другу своими 5'-концевыми частями, которые соответствуют образуемой олигонуклеотидной вставке во фрагмент матричной ДНК. 3'-Концы этих праймеров комплементарны участкам ДНК-матрицы, непосредственно примыкающим к сайту, в который производится вставка олигонуклеотида.

С помощью метода ПЦР удастся легко получать и множественные мутации в конкретных участках ДНК. В этом случае амплификацию мутагенизируемого сегмента ДНК производят в присутствии трех (вместо четырех) дезоксирибонуклеозидтрифосфатов, причем один из них вводят в реакционную смесь в высокой концентрации. Именно этот нуклеотид преимущественно включается в амплифицируемый фрагмент ДНК вместо недостающего нуклеотида, что сопровождается накоплением в ДНК-продукте множественных случайных мутаций в виде соответствующих замен

нуклеотидов. В таких искусственных условиях мутации возникают за счет снижения точности функционирования ДНК-полимеразы.

8.2. Белковая инженерия

После рассмотрения способов получения сайт-специфических мутаций необходимо сделать лишь один шаг, чтобы оказаться лицом к лицу с бурно развивающимся направлением молекулярной генетики, называемым белковой инженерией. Действительно, разработка методов направленного мутагенеза дала возможность не только с высокой точностью модифицировать отдельные белки и изучать их структурно-функциональные взаимоотношения, но и конструировать новые белки, не существовавшие в природе. Впечатляющими результатами применения такого подхода являются гибридные белки, получаемые путем объединения фрагментов и функциональных доменов разных полипептидных цепей с использованием генно-инженерных методов.

Другое перспективное направление белковой инженерии – это конструирование биологически активных пептидов, обладающих фармакологической активностью.

8.2.1. Библиотеки пептидов и эпитопов

В живом организме большинство биологических процессов управляется посредством специфических белок-белковых или белково-нуклеиновых взаимодействий. К таким процессам относятся, например регуляция транскрипции генов под действием различных белковых факторов, взаимодействие белковых лигандов с рецепторами на поверхности клеток, а также специфическое связывание антигенов соответствующими антителами. Понимание молекулярных механизмов взаимодействия белковых лигандов с рецепторами имеет большое фундаментальное и прикладное значение. В частности, разработка новых лекарственных препаратов белковой природы обычно начинается с идентификации исходной последовательности аминокислот, обладающей требуемой биологической активностью (так называемая "основная" (lead) последовательность). Однако пептиды с основной последовательностью аминокислот могут обладать и нежелательными биологическими свойствами: низкой активностью,

токсичностью, малой стабильностью в организме и т.п.

До появления библиотек пептидов улучшение их биологических свойств осуществляли путем последовательного синтеза большого числа аналогов и проверкой их биологической активности, что требовало больших затрат времени и средств. В последние годы появилась возможность с помощью автоматических синтезаторов создавать за короткое время тысячи различных пептидов. Разработанные методы направленного мутагенеза также позволили резко расширить число белков, получаемых одновременно и последовательно тестируемых на биологическую активность. Однако только недавно разработанные подходы к созданию библиотек пептидов привели к получению миллионов последовательностей аминокислот, требуемых для проведения эффективного скрининга с целью выявления среди них пептидов, максимально удовлетворяющих предъявляемым критериям. Такие библиотеки используются для исследования взаимодействия антител с антигенами, получения новых ингибиторов ферментов и антимикробных агентов, конструирования молекул, обладающих требуемой биологической активностью, или придания новых свойств белкам, например антителам.

По способам получения библиотеки пептидов разделяются на три группы. К первой группе можно отнести библиотеки, полученные с использованием химического синтеза пептидов, в которых индивидуальные пептиды иммобилизованы на микроносителях. При таком подходе после присоединения очередных аминокислот в индивидуальных реакционных смесях к пептидам, иммобилизованным на микроносителях, содержимое всех реакционных смесей объединяют и разделяют на новые порции, которые используют на следующей стадии присоединения новых аминокислотных остатков. После проведения ряда таких этапов оказываются синтезированными пептиды, содержащие последовательности использованных в синтезе аминокислот во всевозможных случайных сочетаниях.

Библиотеки пептидов, иммобилизованных на микроносителях, обладают существенным недостатком: они требуют при скрининге использования очищенных рецепторов, находящихся в растворимой форме. В то же время в большинстве случаев при биологических испытаниях, проводящихся для фундаментальных и фармакологических исследований, чаще всего находят применение рецепторы, ассоциированные с мембранами. По второму способу

библиотеки пептидов получают с помощью твердофазного синтеза пептидов, при котором на каждой стадии химического присоединения очередной аминокислоты к растущим пептидным цепям используют эквимольные смеси всех или некоторых аминокислот-предшественников. На конечной стадии синтеза проводят отделение пептидов от носителя, т.е. перевод их в растворимую форму. Третий подход к конструированию библиотек пептидов, к описанию которого мы сейчас переходим, стал реальным именно благодаря развитию методов генной инженерии. Он прекрасно иллюстрирует возможности таких методов и, несомненно, является крупным достижением в их применении. В этой связи рассмотрим более подробно результаты использования библиотек пептидов в исследовании *эпитопов* (антигенных детерминант) белков.

Генно-инженерная технология получения гибридных белков позволила разработать эффективный метод наработки коротких пептидов для анализа их биологической активности. Как и в случае клонотек генов, библиотеки пептидов, полученные генно-инженерными методами, представляют собой большой (часто исчерпывающий) набор коротких пептидов. Два недавно сделанных наблюдения позволяют рассматривать библиотеку пептидов одновременно и в качестве библиотеки эпитопов белков. Во-первых, короткие пептиды могут включать все основные остатки аминокислот, играющие главную роль во взаимодействии с антителами, и они в состоянии имитировать крупные антигенные детерминанты белков. Во-вторых, в большинстве случаев нековалентные связи, образуемые между немногими наиболее важными остатками аминокислот белковых лигандов и их рецепторами, вносят основной вклад в общую энергию взаимодействия лиганд–рецептор. С учетом этого любой пептид можно рассматривать как потенциальный лиганд, гаптен или часть антигенной детерминанты более крупных полипептидов, а любую библиотеку пептидов – как библиотеку эпитопов белков или потенциальных лигандов для соответствующих белковых рецепторов.

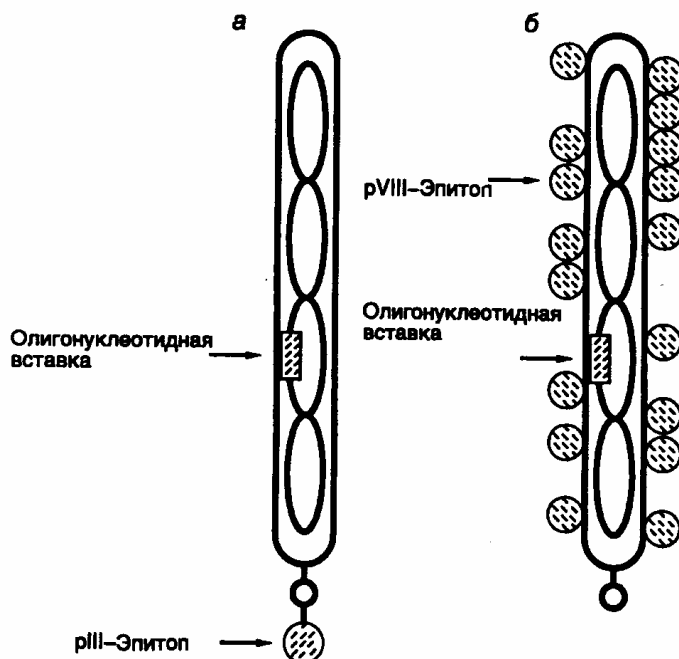


Рис. II.19. Схема экспрессии пептидных эпитопов на поверхности оболочки нитевидных коофагов

Пептидные эпитопы находятся в составе гибридных полипептидных цепей минорного белка рIII (а) или основного белка рVIII вирусной оболочки (б). Стрелки указывают положение кодирующих эпитопы олигонуклеотидных фрагментов в геноме бактериофага, а также положение самих эпитопов. В составе полипептида рIII (а) показана только одна копия эпитопа (на самом деле их число достигает 4–5)

Библиотека пептидов, полученная в результате реализации третьего подхода, в современном виде представляет собой набор десятков или даже сотен миллионов коротких различающихся последовательностей аминокислот, которые экспрессированы на поверхности вирионов бактериофагов в составе их собственных структурных белков. Это становится возможным благодаря введению методами генной инженерии в геном бактериофагов гибридных рекомбинантных генов, кодирующих измененные структурные белки его вирионов. (Данный метод известен под названием *фагового дисплея*.) В результате экспрессии таких генов образуются гибридные белки, на N- или C-концах которых (см. ниже) присутствуют дополнительные последовательности аминокислот. В наиболее хорошо разработанной системе, позволяющей конструировать библиотеки пептидов генно-инженерными методами,

используют небольшой нитевидный колифаг f1 и два его белка: основной и минорный белки оболочки pVIII и pIII. In vivo оба белка синтезируются в виде полипептидных цепей с короткими N-концевыми сигнальными последовательностями, которые отщепляются сигнальной пептидазой во время их созревания после переноса к внутренней части бактериальной мембраны. Зрелые белки встраиваются в оболочку бактериофага в процессе ее сборки. При этом белок pVIII образует основную оболочку бактериофага, тогда как четыре или пять молекул pIII ассоциированы с концевой частью вириона и обеспечивают взаимодействие вирусных частиц с половыми ворсинками клеток *E. coli* (рис. II.19). Генно-инженерными методами пептиды соединяют с белками – непосредственно с их N-концевыми последовательностями или на небольшом от них расстоянии. Концевые последовательности большинства белков являются более гибкими и, как правило, экспонируются на поверхности глобулы, что позволяет получать гибридные рекомбинантные белки без существенного нарушения их основных свойств, а также делает интегрируемые пептиды доступными для распознавания извне. Кроме того, в таком положении и пространственная структура самих пептидов испытывает меньшее влияние белка-носителя. В ходе экспериментов было установлено, что введение чужеродных пептидов в N-концевую часть белка pIII не оказывает существенного влияния на жизнеспособность и инфекционность фаговых частиц, тогда как соединение пептидов длиной >5 аминокислотных остатков с N-концевой частью белка pVIII нарушает сборку вирионов. Последнее затруднение можно преодолеть доставкой к месту сборки вирионов молекул белка pVIII дикого типа, синтез которых направляется соответствующим геном вируса-помощника. В этом случае оболочка бактериофага будет содержать как измененные белки pVIII, так и полипептиды дикого типа от вируса-помощника.

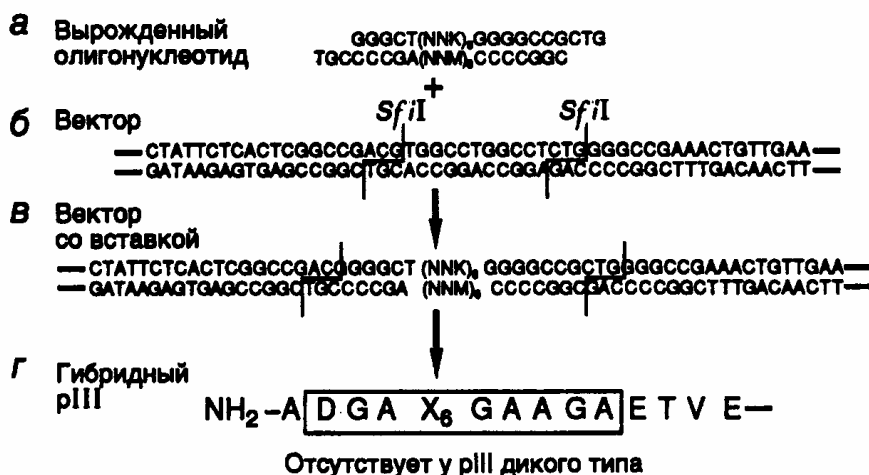


Рис. II.20. Схема конструирования рекомбинантного вирусного генома, содержащего вставки вырожденных олигонуклеотидов, для получения библиотеки эпитопов

Двухцепочечный олигонуклеотид (а), содержащий вырожденные кодоны NNK и те же самые сайты рестрикции в составе линкеров, лигируют с ДНК вектора Fuse5 (б), расщепленного рестриктазой *SfiI*, с образованием рекомбинантного генома (в), который направляет синтез гибридного рекомбинантного белка (г), содержащего на N-конце указанную аминокислотную последовательность

При конструировании библиотеки пептидов прежде всего синтезируют два комплементарных друг другу олигонуклеотида, которые после отжига образуют двухцепочечную молекулу, центральная часть которой кодирует собственно пептиды (рис. II.20,а), а выступающие по концам одноцепочечные участки комплементарны "липким" концам вектора, получающимся под действием соответствующей рестриктазы (см. рис. II.20,б).

Для кодирования аминокислот пептидов используют вырожденные кодоны вида NNK или NNS, которые включают все четыре нуклеотида (N) в первом и втором положениях, G или T (K), а также G или C (S) в третьем положении. При таком подходе информация о всех 20 аминокислотах и одном стоп-кодоне заключена в 32 различных кодонах NNK и NNS, а не в 64, как это имеет место в случае природного генетического кода.

В процессе синтеза вырожденных олигонуклеотидов, кодирующих исследуемые пептиды, на каждой стадии используют индивидуальные нуклеотиды для кодонов инвариантных аминокислот, фланкирующих

вариабельный участок пептида, а также эквимольные смеси нуклеотидов для участков, кодирующих случайные последовательности. Образовавшийся в итоге набор вырожденных олигонуклеотидов далее клонируют в виде одноцепочечных фрагментов в соответствующих сайтах гена белка оболочки бактериофага в составе фагового вектора или фазмиды. Альтернативно для такого набора олигонуклеотидов (химически или с помощью ПЦР) синтезируют комплементарные цепи с включением инозина в вариабельные участки, поскольку его остатки, как известно, спариваются с основаниями С и Т матрицы, что облегчает образование правильных дуплексов между соответствующими олигонуклеотидами. Образующиеся двухцепочечные олигонуклеотиды в случае необходимости обрабатывают соответствующими рестриктазами и клонируют в фаговом векторе. Итоговые рекомбинантные молекулы (см. рис. II.20,в) ДНК вводят в бактериальные клетки, получая $\sim 10^9$ трансформантов на 1 мг рекомбинантной ДНК, образовавшиеся фаговые частицы размножают в бактериях и после очистки исследуют на присутствие рекомбинантных пептидов (см. рис. II.20,г), способных взаимодействовать с исследуемыми рецепторами в белках их вирионов.

Число индивидуальных фаговых клонов в библиотеке является определяющим для ее использования. К примеру, библиотека, заключающая в себе все возможные гексапептиды, должна содержать 64 млн (20^6) разных шестичленных аминокислотных последовательностей, кодируемых ~ 1 млрд (32^6) различных гексакодонов (32 – число кодонов, с помощью которых можно закодировать любую из 20 аминокислот предложенным выше способом, а именно с использованием кодонов NNK или NNS). Для решения такой задачи должны быть получены очень большие библиотеки, содержащие, по крайней мере, $2 \cdot 10^8 - 3 \cdot 10^8$ индивидуальных, независимых клонов, а величина 10^9 в настоящее время является верхним пределом для числа индивидуальных клонов библиотеки, которую еще можно практически использовать.

Исходя из этого, можно заключить, что максимальная длина пептидов, включающих в себя все возможные сочетания 20 аминокислот, с которыми возможно работать с помощью библиотек пептидов, составляет 6 аминокислотных остатков. Тем не менее, следует иметь в виду, что библиотека 15-членных пептидов того же размера ($2-3 \cdot 10^8$ клонов) будет содержать больше разнообразных гексапептидов, чем рассмотренная выше библиотека 6-

членных пептидов. Кроме того, поскольку лишь ограниченное число аминокислотных остатков в пептиде действительно определяет его биологическую активность, библиотека 15-членных пептидов может оказаться представительнее библиотеки более коротких пептидов с тем же числом клонов.

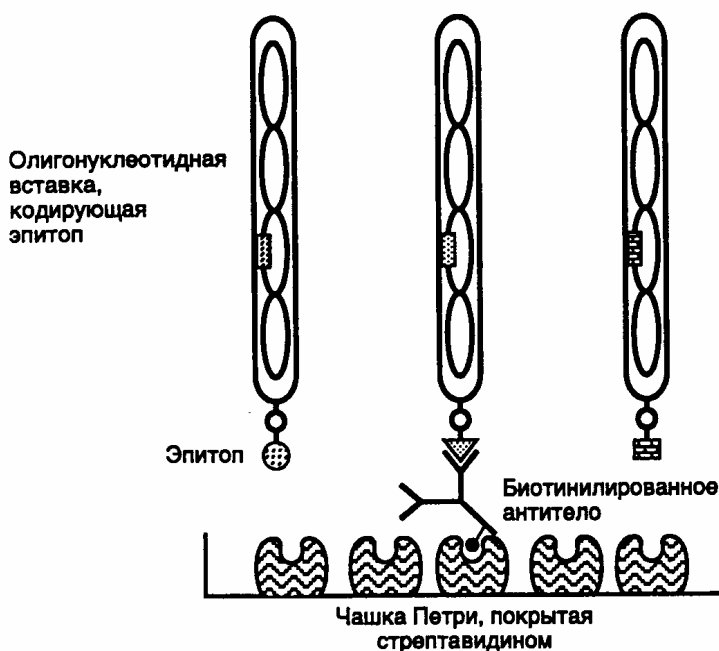


Рис. II.21. Схема отбора фаговых частиц, обладающих требуемыми эпитопами

Показаны три рекомбинантных фаговых частицы, экспрессирующие разные эпитопы в составе pIII. Только эпитоп центральной фаговой частицы распознается молекулой биотинилированного антитела, иммобилизованного на чашке Петри с помощью стрептавидина и использованного для скрининга библиотеки

Для того чтобы выделить из библиотеки пептиды с искомой биологической активностью, применяют различные методы скрининга. В частности, для выделения пептидов, имитирующих определенные эпитопы, используют биотинилированные моноклональные антитела соответствующей специфичности, которые иммобилизуют на твердой подложке с помощью стрептавидина (рис. II.21). Фаговые частицы, экспрессирующие на своей поверхности соответствующие эпитопы, взаимодействуют с антителами и задерживаются подложкой, тогда как другие рекомбинантные фаговые частицы

удаляются в процессе промывания. Задержанные на подложке фаговые частицы далее элюируют кислотой, индивидуальные клоны дополнительно размножают в бактериальных клетках и экспрессированные на них эпитопы исследуют по различным критериям. Наличие идентичных или сходных последовательностей нуклеотидов среди клонированных последовательностей свидетельствует о специфичности процесса очистки. Индивидуальные клоны затем охарактеризовывают другими, в частности иммуноферментными методами. На заключительной стадии исследования осуществляют синтез выделенных пептидов и их всестороннее изучение в очищенном состоянии.

В настоящее время имеются данные о некоторых работах, проведенных с использованием пептидных библиотек. В одном из таких исследований из библиотеки были выделены пептиды, последовательность аминокислот в которых резко отличалась от последовательности аминокислот истинного эпитопа исследуемого антигена. Тем не менее, такой пептид прочно связывался со специфическими антителами и конкурировал за связывание с природным антигеном. Это позволило сделать вывод о возможности существования *мимотопов* – коротких пептидов, имитирующих природные эпитопы, последовательности аминокислот которых существенно различаются между собой. Удалось установить канонические последовательности аминокислот пептидов, имитирующих эпитопы природных белков, а среди них идентифицировать аминокислотные остатки, играющие ключевую роль во взаимодействии антиген–антитело.

Одним из многообещающих приложений библиотек пептидов является идентификация пептидных лигандов, имитирующих "структурные" эпитопы, образующиеся на поверхности белковых глобул в результате сворачивания их полипептидных цепей, что сопровождается пространственным сближением аминокислотных остатков, расположенных в полипептидной цепи на значительном расстоянии друг от друга. С помощью пептидных библиотек возможна идентификация пептидных аналогов различных эпитопов небелковой природы. По-видимому, в ближайшем будущем возможно использование пептидных библиотек для получения новых лекарственных препаратов, создания диагностических средств и производства эффективных вакцин. В области конструирования новых лекарственных препаратов усилия исследователей могли бы быть направлены на создание пептидных лигандов,

специфически взаимодействующих с рецепторами, представляющими медико-биологический интерес. Знание структуры таких лигандов позволило бы упростить получение на этой основе лекарственных препаратов небелковой природы.

Библиотеки пептидов и эпитопов найдут свое применение и в исследованиях механизмов гуморального иммунного ответа, а также заболеваний иммунной системы. В частности, большинство аутоиммунных заболеваний сопровождается образованием аутоантител против антигенов собственного организма. Эти антитела во многих случаях служат специфическими маркерами того или иного аутоиммунного заболевания. С использованием библиотеки эпитопов, в принципе, можно получить пептидные маркеры, с помощью которых было бы возможно следить за специфичностью аутоантител во время развития патологического процесса как в индивидуальном организме, так и в группе пациентов и, кроме того, определять специфичность аутоантител при заболеваниях неизвестной этиологии.

Библиотеки пептидов и эпитопов потенциально могут быть использованы также для скрининга иммунных сывороток с целью выявления пептидов, специфически взаимодействующих с защитными антителами. Такие пептиды будут имитировать антигенные детерминанты патогенных организмов и служить мишенями для защитных антител организма. Это позволит использовать подобные пептиды для вакцинации пациентов, у которых отсутствуют антитела против соответствующих патогенов. Изучение эпитопов с помощью библиотек пептидов является частным случаем одного из многочисленных направлений их использования в прикладных и фундаментальных исследованиях взаимодействия лигандов и рецепторов. Дальнейшее усовершенствование этого подхода должно способствовать созданию новых лекарственных препаратов на основе коротких пептидов и быть полезным в фундаментальных исследованиях механизмов белок-белковых взаимодействий.

8.2.2. Белки-репортеры в гибридных белках

В рассмотренных выше библиотеках пептидов последние ковалентно связаны с белком-носителем. В таком виде они являются одними из представителей гибридных белков, получаемых методами генной инженерии.

В другом случае гибридные белки применяют для получения высокого уровня экспрессии коротких пептидов в бактериальных клетках благодаря стабилизации этих пептидов в составе гибридных белков. Часто гибридные белки используют для идентификации и очистки трудноопределяемых рекомбинантных белков. Например, присоединив к С-концу исследуемого белка в качестве белка-репортера β -галактозидазу, можно производить очистку рекомбинантного белка по активности β -галактозидазы, определяя ее антигенные детерминанты иммунохимическими методами. Соединяя фрагменты ДНК, содержащие открытые рамки считывания (ОРС), с генами белков-репортеров, можно очистить такие гибридные белки по активности белка-репортера и использовать их для иммунизации лабораторных животных. Полученные антитела далее применяют для очистки нативного белка, в состав которого входит рекомбинантный полипептид, кодируемый ОРС, и тем самым идентифицируют клонированный фрагмент гена.

С помощью гибридных белков решают и обратную задачу клонирования неизвестного гена, к белковому продукту которого имеются антитела. В таком случае конструируют клонотеку последовательностей нуклеотидов, представляющих ОРС неизвестных генов, в векторах, которые позволяют соединять клонируемую ОРС в одной рамке считывания с геном-репортером. Образующиеся в результате экспрессии этих рекомбинантных генов гибридные белки идентифицируются с помощью антител иммуноферментными методами. Гибридные гены, объединяющие секретируемые белки и белки-репортеры, дают возможность по-новому исследовать механизмы секреции, а также локализацию и перемещение в тканях секретируемых белков.

8.2.3. Гибридные токсины

Серия работ И. Пастана с сотрудниками по конструированию гибридных токсинов направленного действия прекрасно иллюстрирует возможности белковой инженерии в части комбинирования различных функциональных доменов белков для достижения конкретных биологических эффектов.

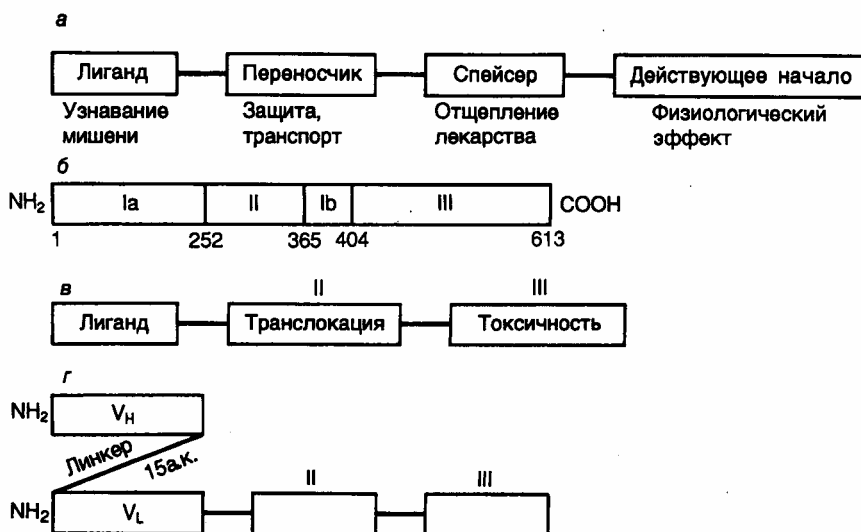


Рис. II.22. Лекарственные препараты направленного действия на основе гибридных токсинов

а – обобщенная схема структуры лекарственного препарата направленного действия; **б** – строение псевдомонадного токсина (цифрами обозначено положение аминокислотных остатков); **в** – строение гибридного токсина; **г** – гибридный токсин на основе моноклональных антител

Идеальное лекарственное средство строго специфического избирательного действия должно обладать, по крайней мере, следующими структурно-функциональными особенностями (рис. II.22,а). Такой лекарственный препарат должен заключать в себе действующее начало для достижения физиологического эффекта и лиганд, распознающий рецептор на поверхности клеток-мишеней. Кроме того, в нем должны быть структурные элементы, распознаваемые системой транспорта организма, для доставки лекарства к клеткам-мишеням, а также спейсерный участок, необходимый для отделения действующего начала от остальных функциональных частей препарата после его доставки по адресу. Именно такая идеальная схема реализуется в природном экзотоксине *Pseudomonas aeruginosa*. Экзотоксин *P. aeruginosa* представляет собой белок, состоящий из одной полипептидной цепи длиной в 613 аминокислот, которая организована в три функциональных домена (см. рис. II.22,б). N-Концевой домен Ia (аминокислотные остатки 1–252) необходим для взаимодействия с поверхностью клеток-мишеней (прототип лиганда идеального лекарства направленного действия). Функции домена Ib

(аминокислотные остатки 365–404) в настоящее время неизвестны. Домен II (аминокислотные остатки 253–364) обеспечивает эффективный перенос токсина в цитозоль клеток (система транспорта лекарства), а домен III (аминокислотные остатки 405–613) осуществляет ADP-рибозилирование фактора элонгации трансляции EF2, что приводит к подавлению трансляции и гибели клеток-мишеней. Таким образом, для оказания цитотоксического действия экзотоксину А необходимо с помощью домена Ia распознать рецепторы на поверхности клеток, проникнуть в клетку с помощью эндоцитоза, опосредованного рецепторами, и быть транслоцированным через внутреннюю мембрану в цитозоль, где локализуется фактор EF2. Основная идея в создании токсинов направленного действия заключалась в том, чтобы заменить домен Ia на какой-либо иной пептидный лиганд, взаимодействующий с другой группой рецепторов на поверхности клеток, и тем самым изменить специфичность действия токсина в отношении самих клеток (см. рис. II.22, в).

Было установлено, что удаление домена Ia генно-инженерными методами резко (в сотни и тысячи раз) снижает токсичность такого укороченного белка как в отношении клеток различных линий, так и *in vivo*. Присоединение к С-концевой части укороченного полипептида молекулы интерлейкина 2 человека осуществляли путем объединения структурных частей соответствующих генов в экспрессирующем векторе. Очищенный гибридный токсин оказался чрезвычайно токсичным в отношении клеток, несущих на своей поверхности рецепторы интерлейкина 2, и не действовал на клетки, у которых эти рецепторы отсутствовали и которые погибали под действием природного токсина. *Интернализация* (транслокация внутрь клеток) гибридного токсина была опосредована субъединицами p55 и p70 рецептора интерлейкина 2. Таким образом, в результате действия гибридного токсина на популяцию клеток, часть из которых экспрессирует на своей поверхности рецепторы интерлейкина 2, происходит избирательная гибель именно этих клеток.

В организме большинство покоящихся Т-клеток и Т-клеток памяти не экспрессируют на своей поверхности высокоаффинных рецепторов интерлейкина 2, тогда как Т-клетки, стимулированные аллоантигенами, содержат такие рецепторы. Поэтому внутрибрюшинное введение гибридного токсина крысам с экспериментальным артритом – заболеванием,

обусловленным патологической активацией Т-клеток, снижало симптомы заболевания. Гибридный токсин существенно уменьшал у мышей и реакции отторжения трансплантата.

Вслед за этими пионерскими работами последовала целая серия исследований, направленных на создание аналогичных систем адресной доставки различных цитотоксических полипептидов. В процессе дальнейшего усовершенствования системы адресной доставки псевдомонадного токсина с использованием интерлейкина 2 в качестве лиганда отказались от полного удаления адресного домена токсина и ограничились его инактивацией путем введения в ген токсина четырех сайт-специфических мутаций. Молекулы такого гибридного токсина оказались в 10–100 раз более эффективными цитотоксическими агентами против клеток человека и обезьян, экспрессирующих на своей поверхности рецепторы для интерлейкина 2, а также обладали значительно большим временем полужизни в крови мышей *in vivo* по сравнению с ранее полученной конструкцией.

На основе псевдомонадного токсина были созданы гибридные токсины, содержащие в качестве лигандов полипептидные цепи интерлейкина 4, интерлейкина 6, трансформирующего фактора роста типа α и инсулиноподобного фактора роста I. Для всех этих гибридных белков была показана высокоспецифическая цитотоксичность в отношении опухолевых клеток (включая клетки миеломы человека), обладающих соответствующими рецепторами. Использование в гибридном токсине в качестве лиганда части полипептидной цепи CD4 – гликопротеина поверхности Т-клеток, который является рецептором вируса ВИЧ и взаимодействует с его гликопротеином gp120, позволило избирательно поражать Т-клетки, зараженные вирусом ВИЧ и экспрессирующие на своей поверхности вирусный белок gp120.

Тот же принцип подавления инфекции, вызванной вирусами ВИЧ, растворимыми рецепторами CD4 был использован при конструировании гибридных белков, объединяющих части полипептидных цепей CD4 с константными частями тяжелых или легких цепей иммуноглобулинов человека. При этом в процессе объединения генов были удалены последовательности нуклеотидов, кодирующие трансмембранный и цитоплазматический домены CD4, а также варибельную часть полипептидных цепей иммуноглобулинов. Образующиеся гибридные молекулы, названные *иммуноадгезинами*, за счет

константной части молекулы иммуноглобулина приобретали повышенную стабильность в организме и, кроме того, сохраняли специфические свойства, опосредуемые константными частями иммуноглобулинов: связывание Fc-рецептора и белка А, способность к фиксации комплемента и перенос через плацентарный барьер. Совокупность всех этих свойств давала возможность иммуноадгезинам эффективно прерывать инфекцию Т-клеток вирусом ВИЧ-1, блокируя как сам вирус, так и зараженные им клетки, экспрессирующие на своей поверхности вирусный антиген gp120.

Дальнейшее усовершенствование генно-инженерных конструкций на основе псевдомонадного экзотоксина А произошло после того, как в качестве адресной части гибридного токсина стали использовать вариабельные домены моноклональных антител к компоненту р55 рецептора интерлейкина 2 человека. В этом рекомбинантном белке с помощью 15-звенного пептидного линкера аминокислот соединяли вариабельный домен тяжелой цепи этого иммуноглобулина с вариабельным доменом его легкой цепи, а С-конец легкой цепи – с N-концом укороченного псевдомонадного токсина (см. рис. II.22,а). Такие молекулы гибридного токсина также оказались высокоспецифичными цитотоксическими агентами по отношению к лейкозным клеткам человека, экспрессирующим на своей поверхности рецепторы интерлейкина 2.

Разработанный подход продемонстрировал возможность использования специфических антител в качестве адресных частей гибридных токсинов. Это дает в руки исследователей универсальный способ адресной доставки токсинов, который в будущем позволит оказывать цитотоксическое действие на любые группы клеток, экспрессирующих на своей поверхности специфические антигены, т.е. значительно расширить количество мишеней для химиотерапевтических воздействий с использованием рекомбинантных белков.

Помимо псевдомонадного экзотоксина А в качестве действующего начала в гибридных токсинах успешно применяли дифтерийный токсин, фактор некроза опухолей и А-цепь рицина. Поскольку А-белок избирательно взаимодействует с константными (Fc) частями иммуноглобулинов класса G многих млекопитающих, такой гибридный токсин в паре с иммуноглобулином, полученным против какого-либо антигена на поверхности клеток, избирательно связывается с этими клетками и убивает их. Такие иммунотоксины являются еще одним потенциальным противоопухолевым агентом и могут быть

использованы против клеток, экспрессирующих на своей поверхности специфические антигены.

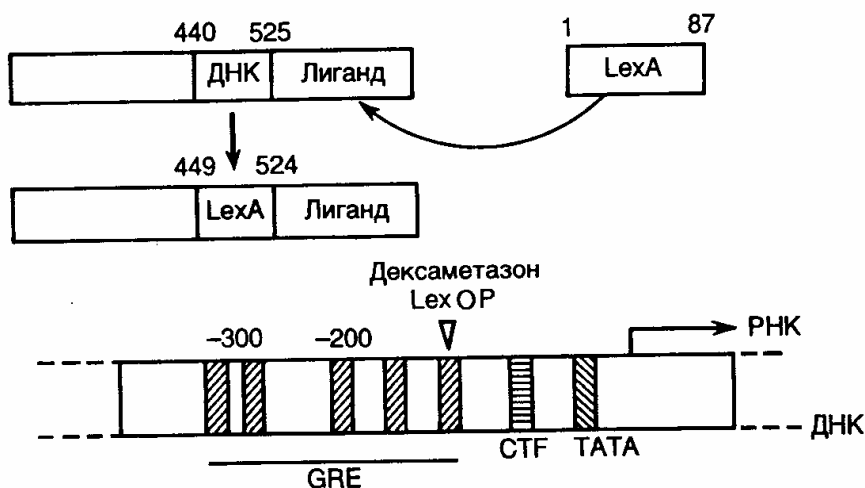


Рис. II.23. Использование гибридного белка для регуляции экспрессии гена

Методы генной инженерии открывают безграничные возможности конструирования новых белков путем объединения в разных комбинациях различных функциональных доменов полипептидных цепей. Получение гибридных токсинов направленного действия иллюстрирует возможности такого подхода в белковой инженерии. В качестве последней иллюстрации возможностей этой группы методов рассмотрим гибридный белок как новый регулятор активности генов. При конструировании такого белка методами генной инженерии был заменен ДНК-связывающий домен в рецепторе глюкокортикоидных гормонов на соответствующий домен LexA-репрессора *E. coli* (рис. II.23).

Введение операторной последовательности гена *lexA* в область промотора глобинового гена (или других генов) приводило к активации промотора под действием гибридного белка в присутствии дексаметазона – синтетического гормона, взаимодействующего с рецептором глюкокортикоидов. Таким образом, в новом генетическом окружении последовательность нуклеотидов оператора гена *lexA* *E. coli* функционировала в качестве энхансера транскрипции в присутствии гибридного белка-активатора, узнающего эту последовательность. Результаты работы демонстрируют возможность

создания новых белков – регуляторов активности генов путем комбинирования известных функциональных доменов.

Развитие белковой инженерии во многом сдерживается недостатком знания о структурно-функциональных взаимоотношениях в белках, что обусловлено сложностью объекта исследования. Многочисленные работы, направленные на изучение таких связей, как правило, носят эмпирический характер и завершаются локализацией аминокислот, существенных для функционирования активных центров ферментов. Поэтому основная задача белковой инженерии – по известной последовательности аминокислотных остатков получить белок с заданными свойствами – в настоящее время еще далека от своего разрешения. Тем не менее, уже сейчас иногда удается целенаправленно изменять некоторые свойства существующих ферментов путем замен небольшого числа аминокислотных остатков их полипептидных цепей с помощью направленного мутагенеза.

8.2.4. Подходы к созданию новых ферментов

Подавляющее большинство исследований, в которых методы белковой инженерии используют для замен отдельных аминокислотных остатков в полипептидных цепях белков, заканчиваются получением мутантных производных, у которых отсутствует, резко ослаблена или остается неизменной исходная ферментативная активность. В таких опытах обычно локализуют отдельные остатки аминокислот, участвующих в формировании активных центров ферментов. Однако в ряде случаев методами направленного мутагенеза удается целенаправленно и кардинально изменять свойства белков. В опытах такого рода можно повысить термостабильность исследуемых белков, усилить их устойчивость по отношению к окислению, протеолизу и другим неблагоприятным воздействиям, а также изменить у исходных ферментов субстратную специфичность.

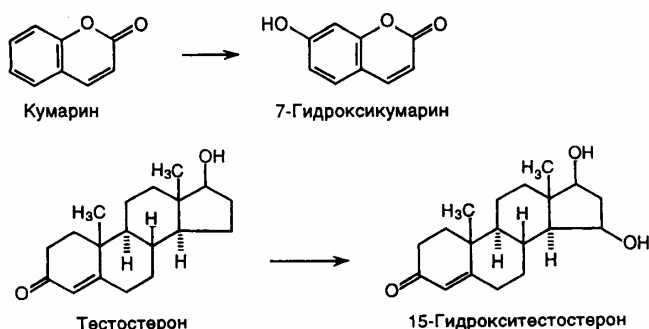
Серия работ была проведена на субтилине – протеолитическом ферменте бактерий рода *Bacillus*. В активном центре субтилина *B. lentus* в положении 222 находится метионин. Окисление Met-222 сопровождается образованием соответствующего сульфоксида и полностью инактивирует фермент. Методом направленного мутагенеза произвели замену Met-222→Cys

и остатки цистеина модифицировали тиаалкилирующим агентом. Это делало фермент высокоустойчивым к окислению и сопровождалось лишь незначительной потерей его активности. Новые свойства модифицированного субтилизина расширяют возможности его применения в биотехнологии. Путем замены отдельных аминокислотных остатков субтилизина в других работах удавалось изменять его субстратную специфичность, повышать удельную активность, изменять оптимум pH и увеличивать устойчивость по отношению к щелочи.

Эти работы иллюстрируют возможности данного направления исследований в белковой инженерии, направленных на создание высокотехнологичных ферментов, способных длительное время функционировать в жестких условиях ферментации на промышленных установках. На еще одно направление развития исследований в белковой инженерии указывают работы по направленному изменению субстратной специфичности ферментов путем замены отдельных аминокислотных остатков в их полипептидных цепях. Например, замена Met-73 у ингибитора субтилизина *Streptomyces lividans* на Lys или Arg приводила к появлению у него способности ингибировать трипсин. При этом у мутантного ингибитора с заменой Met-73→Lys появлялась также способность ингибировать лизилэндопептидазу, а ингибитор с заменой Met-73 на Tyr или Trp подавлял активность α -трипсиногена. Исследования продемонстрировали функциональную значимость остатка Met-73 в так называемом PI-сайте участка полипептидной цепи, контактирующего с ингибируемыми ферментами, и позволили создать модель такого белок-белкового взаимодействия. Кроме того, эти работы являются основой для конструирования новых ингибиторов сериновых протеиназ, а также более глубокого изучения механизма их ингибирующего действия.

Такой же подход с целенаправленными заменами отдельных аминокислотных остатков был использован для изменения субстратной специфичности трипсина. В этом случае предварительный теоретический анализ структурных особенностей активного центра трипсина и особенно его субстрат-связывающего участка позволил предсказать аминокислотные остатки, существенные для взаимодействия субстрата с ферментом и обеспечивающие специфичность его действия. Проведенная на основании

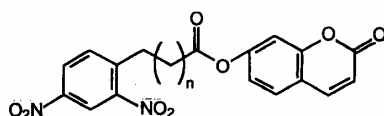
такого анализа замена Gly-226 на Ala резко (в 100 раз) уменьшила способность мутантного трипсина гидролизовать субстраты, содержащие Arg, и в значительно меньшей степени повлияла на его активность в отношении лизиновых субстратов. Обратную, хотя и менее ярко выраженную, картину наблюдали при замене Gly-216 → Ala. Проведенная группой И. Сигала замена Ser-70 в активном центре β -лактамазы на Cys также изменяла ее субстратную специфичность. Константа связывания субстрата – пенициллина – мутантным ферментом оставалась неизменной, тогда как скорость его расщепления уменьшалась в 50–100 раз. В то же время мутантная β -лактамаза была более активна, чем фермент дикого типа по отношению к цефалоспорином – аналогам пенициллина, так называемым антибиотикам третьего поколения. Возможность изменения субстратной специфичности путем замены отдельных аминокислот была продемонстрирована недавно и для цитохромов группы P-450. Мышиные цитохромы P-450_{15 α} и P-450_{coh} экспрессируются в клетках печени самок и почек самцов соответственно. Первый из них катализирует 15 α -гидроксилирование Δ^4 -3-кетостероидов типа тестостерона, тогда как второй – гидроксилирование кумарина в положении 7. Несмотря на свою разную субстратную специфичность оба фермента различаются лишь по 11 (из 494) аминокислотным остаткам. Осуществляя последовательные замены этих 11 аминокислот, установили, что замена единственной аминокислоты Phe-209→Leu в полипептидной цепи цитохрома P-450_{coh} приводит к появлению у него способности гидроксилировать стероиды, т.е. активности цитохрома P-450_{15 α} , и к потере способности использовать кумарин в качестве субстрата.



Эти далеко не исчерпывающие примеры демонстрируют возможности

одного из двух основных подходов к созданию методами белковой инженерии белков с новыми каталитическими активностями. Такой подход заключается в изменении существующих активных центров ферментов путем замен отдельных аминокислотных остатков, которые участвуют в формировании активных центров. Принципиально другой подход к получению ферментов новой специфичности заключается в модификации методами белковой инженерии биологических рецепторов путем введения каталитических функциональных групп в центры связывания лигандов-субстратов. Такой подход был, в частности, реализован при конструировании методом направленного мутагенеза каталитических антител – *абзимов*.

Была поставлена задача методами белковой инженерии на основе иммуноглобулинов, специфически взаимодействующих с динитрофенильными группами антигенов, получить абзимы, гидролизующие эфиры динитрофенола следующей структуры:



Исходя из того, что имидазол действует в качестве нуклеофильного катализатора при гидролизе карбоксиэфиров в водных растворах, предполагали, что введение остатка гистидина в центр связывания динитрофенильных лигандов иммуноглобулина может придать ему способность осуществлять гидролиз сложных эфиров динитрофенола. В качестве исходного белка при конструировании абзима использовали V_L - и V_H -фрагменты тяжелой и легкой цепей иммуноглобулина, специфически взаимодействующего с динитрофенильными группами антигенов. V_L - и V_H -фрагменты представляют собой переменные области иммуноглобулина, которые взаимодействуют с антигеном. Фрагмент ДНК, кодирующий V_L -область легкой цепи, синтезировали *in vitro* и экспрессировали в клетках *E. coli*. При этом в процессе синтеза кодон, кодирующий Туг-34, был изменен на кодон His, так как было известно, что именно Туг-34 контактирует с молекулами динитрофенола в комплексах антиген–антитело. Очищенный рекомбинантный V_L -фрагмент объединяли с V_H -

фрагментом тяжелой цепи природного иммуноглобулина, взаимодействующего с динитрофенолом, что приводило к образованию так называемого Fv-фрагмента.

Оказалось, что такой мутантный Fv-фрагмент способен катализировать гидролиз кумаринового эфира 5-(2,4-динитрофенил)аминопентановой кислоты (см. выше структурную формулу, $n=4$) в 90 000 раз более эффективно, чем 4-метилимидазол, и начальная скорость гидролиза была в 45 раз выше, чем у Fv-фрагмента дикого типа. При этом мутантный Fv-фрагмент был неактивен в отношении аминпропанового ($n=2$) и амингексанового ($n=5$) аналогов субстрата. Таким образом, с помощью замены единственной аминокислоты в антигенсвязывающем центре иммуноглобулина удалось создать на его основе высокоспецифичный гидролитический фермент, а сам процесс его конструирования был основан на знании молекулярных механизмов катализа и взаимодействия антигенов с антителами.

В настоящее время реализованы оба подхода к созданию искусственных ферментов с новыми субстратными специфичностями, основанные на изменении отдельных аминокислотных остатков в субстрат- или лигандсвязывающих центрах ферментов или рецепторов. По мере углубления знаний о механизмах действия ферментов и закономерностей реализации потенций первичной структуры полипептидных цепей в их пространственной структуре появится возможность конструирования абсолютно новых белков с заданной ферментативной активностью. С этого момента значительно расширится поле деятельности белковой инженерии, которая вместе с генной инженерией сможет решать задачи, связанные с целенаправленным изменением уже известных и созданием новых метаболических путей в клетках живых организмов.

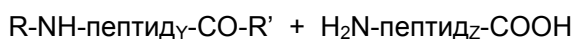
8.2.5. Субтилигаза в лигировании пептидов

В заключение рассмотрим еще одно неожиданное направление белковой инженерии, четко обозначившееся в самое последнее время. Во всех вышеупомянутых подходах конструирование белков с новыми свойствами основано на использовании молекулярно-генетических методов направленного мутагенеза. К сожалению, все эти методы обладают одним существенным недостатком: с их помощью невозможно осуществить встраивание в

исследуемые полипептидные цепи многих неприродных аналогов аминокислот или других химических соединений, что весьма существенно ограничивает возможности структурно-функциональных исследований белков. Такие модификации можно было бы вносить в процессе химического синтеза пептидов, однако в этом случае невозможно получать пептиды большой длины (как правило, >40 остатков аминокислот) из-за накопления побочных продуктов синтеза, что затрудняет последующую очистку основного продукта и снижает выход пептидов.

Недавно был разработан альтернативный подход к химическому синтезу длинных пептидов, заключающийся в ферментативном лигировании предварительно синтезированных химическими методами коротких пептидов друг с другом. Этот подход стал возможен благодаря созданию с помощью направленного мутагенеза аналога субтилизина, так называемого субтилизина BPN', катализирующего образование пептидных связей в водных растворах. Такой мутантный фермент, получивший название *субтилигазы*, содержит в своей полипептидной цепи две мутационные замены: Ser заменен на остаток Cys, а Pro – на Ala. Первая замена в активном центре субтилизина резко повышает его способность к аминолизу (т.е. образованию пептидной связи) по сравнению с гидролитической активностью при использовании в качестве субстратов тетрапептидных эфиров. Вторая мутация изменяет конформацию мутантного активного центра, компенсируя стерические эффекты, вызванные введением остатка Cys, что еще более улучшает каталитические свойства фермента.

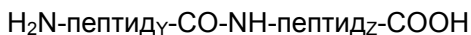
Д. Джексоном с соавторами (1994 г.) разработана стратегия синтеза крупных полипептидных цепей с использованием субтилигазы, реализованная в синтезе модифицированной рибонуклеазы А, обладающей ферментативной активностью. На первом этапе синтеза был получен полностью незащищенный пептид, представляющий собой С-концевой фрагмент РНКазы А, и он был использован в дальнейшем в качестве акцепторной молекулы (см. схему).



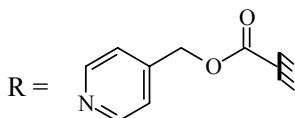
↓ 1. Субтилигаза



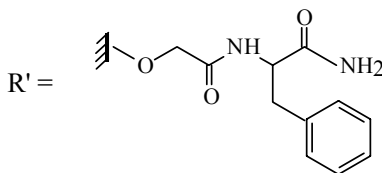
↓ 2. Zn/CH₃COOH



R-NH-пептид_X-CO-R' ↓ 3. Повторение этапов 1 и 2



Изоникотиноил



Гликолат-фенилаланиламид (glc-F-NH₂);

X, Y, Z – различные пептиды

Следующий N-концевой донорный фрагмент полипептидной цепи этерифицирован по C-концу гликолат-фенилаланиламидом (glc-F-NH₂). Полученный эфир эффективно ацилируется субтилигазой, что отражает особенности ее субстратной специфичности: фермент предпочитает использовать в качестве субстрата эфиры, в которых освобождается группа glc-F-NH₂. В отличие от этого донорный пептид содержит на N-конце дополнительную изоникотиноильную защитную группу, что предотвращает его лигирование самого на себя. Такая группа вводится на последней стадии твердофазного синтеза пептидов и проявляет устойчивость к разбавленной плавиковой кислоте (HF), которая использовалась для снятия защиты с боковых цепей пептидов и отщепления самих пептидов от твердого носителя. После каждого лигирования пептидов изоникотиноильную группу удаляли в условиях мягкого восстановления (в присутствии цинка и уксусной кислоты) для освобождения NH₂-группы – необходимой участницы очередного цикла лигирования.

Исходя из субстратной специфичности субтилигазы (фермент эффективно использует крупные гидрофобные донорные субстраты и

малозффективен с отрицательно заряженными остатками аминокислот или остатками пролина, попадающими в его активный центр), карту полипептидной цепи РНКазы А разделили на шесть фрагментов. Эти фрагменты были синтезированы и лигированы друг с другом в соответствии с вышеприведенной схемой с выходом, достигающим 70% на каждом этапе. В результате была получена полноразмерная полипептидная цепь РНКазы А (124 аминокислотных остатка), которая после спонтанного фолдинга приобрела искомую ферментативную активность. В ходе синтеза отдельных пептидов исследователи заменили два остатка His, находящихся в активном центре фермента (положения 12 и 119), на остатки 4-фторгистидина. Поскольку это производное His обладает $pK_a = 3,5$ (pK_a His = 6,8), у искусственной РНКазы А наблюдали резкий сдвиг в оптимуме pH реакции гидролиза субстрата без заметных изменений $k_{кат}$, что было особенно неожиданным результатом, позволившим сделать интересные выводы о механизме ферментативной реакции, осуществляемой РНКазой А. Позднее субтилизин был успешно использован для синтеза модифицированного гормона роста человека и циклических пептидов.

Значение результатов, полученных в ходе направленного изменения субстратной специфичности субтилизина, превратившегося в субтилигазу, выходит далеко за рамки чисто прикладного использования нового искусственного фермента. Это и многие другие подобные исследования, осуществленные методами белковой инженерии, однозначно указывают на возможность изменения субстратной специфичности известных ферментов путем замены в них одного или нескольких критических аминокислотных остатков. Такого рода результаты, примеры которых уже были приведены выше и более подробно будут рассмотрены в следующем разделе, приподнимают завесу над механизмами молекулярной эволюции метаболических путей живого организма, лежащих в основе его существования. История развития молекулярной биологии и генетики показывает, что практически любой неизвестный в природе механизм, разработанный и реализованный исследователем в физиологических условиях в пробирке, уже используется *in vivo* в существующих генетических системах. Однажды на семинаре, когда кто-то пытался выдать банальность за новую информацию, Р.Б. Хесин сказал: "Это новость только для того, кто не читает научной литературы". Мы записаны

в библиотеку Природы, однако, несмотря на все усилия, в настоящее время в состоянии читать на разных языках только букварь, написанный под ее диктовку. Вся область белковой инженерии, которая, на первый взгляд, кажется блестящим порождением лишь одного интеллекта человека, не есть исключение, поскольку ферменты с новой субстратной специфичностью могут быть результатом точковых мутаций, непрерывно происходящих в геноме любого живого организма. А следовательно, Природа непрерывно конструирует новые белки и отбирает лучшие варианты с самого момента возникновения жизни на Земле. И этот процесс ежеминутно происходит в биосфере, в том числе и в онтогенезе эукариот.

8.3. Концепция ксенобиоза

Успехи белковой инженерии, демонстрирующие возможность изменения субстратной специфичности ферментов путем замены одной или нескольких аминокислот с помощью направленного мутагенеза, наводят на многочисленные размышления. В частности, возникает вопрос: почему изменения субстратной специфичности ферментов являются редким событием, а мутации в конкретном гене, как правило, сопровождаются ослаблением или потерей ферментативной активности вообще? В силу ограниченности наших знаний о структурно-функциональных взаимоотношениях в белках ответ на такой вопрос кажется очевидным: одиночные замены аминокислот в активном центре фермента или его окрестностях приводят к конформационным или иным изменениям полипептидной цепи, делающим активный центр нефункциональным в отношении своего природного субстрата.

Вероятно, на самом деле последствия таких мутационных замен аминокислот могут быть более значительными. Начнем с того, что внимание исследователей, изучающих свойства мутантных ферментов, как правило, целиком направлено на изучение влияния мутаций на его основную активность. При этом не учитывается возможность появления новой ферментативной активности, которая остается незамеченной просто потому, что исходно неизвестно, появления какой активности следует ожидать. А между тем при ближайшем рассмотрении такие последствия почти любой миссенс-мутации (точковой мутации, которая приводит к замене одного осмысленного кодона на другой) в структурной части гена, кодирующего полипептидную цепь, кажутся

весьма вероятными. Действительно, в природе, по-видимому, не существует ферментов с абсолютной субстратной специфичностью, и полифункциональность является изначальным и фундаментальным свойством сложных полипептидов. Какой бы узкой ни была субстратная специфичность ферментов по отношению к природным субстратам, для них всегда можно синтезировать искусственные субстраты, расширяющие их субстратную специфичность. Полифункциональность описана для многих природных белков. Например, кристаллины (одни из основных белков хрусталика глаза позвоночных) у млекопитающих являются одновременно малым белком теплового шока неизвестной функции, у птиц и рептилий обладают активностью аргининосукцинатлиазы, другая же форма (δ) у рыб и рептилий обладает лактатдегидрогеназной активностью и является ДНК-связывающим белком. Изоформа кристаллинов τ , характерная для рыб, птиц и рептилий, является одновременно α -энолазой и белком теплового шока. У гемоглобина – основного переносчика кислорода и углекислого газа животных, обнаружены анилингидроксилазная, каталазная и пероксидазная активности. Все вышеперечисленные активности гемоглобина, по-видимому, не имеют физиологического значения и могли возникнуть именно как побочные активности в процессе эволюционирования основной функции этого белка.

Еще более яркий пример такого рода демонстрируют абзимы (антитела, обладающие ферментативной активностью). Помимо хорошо известной гидролитической активности абзимов в отношении различных эфиров и кокаина описана их способность осуществлять синтез амидов, реакции стереоспецифической циклизации, перегруппировку Клайзена, приводящую к образованию префеновой кислоты из хоризмовой, а также ряд других реакций. Сколько же потенциальных ферментативных активностей заключено во всей популяции антител одного индивидуума, репертуар которых практически безграничен благодаря запрограммированному процессу соматического мутагенеза?!

Исходя из этого можно ожидать появления в организме человека и животных антител, обладающих самыми неожиданными ферментативными активностями, гены большей части которых элиминируются в процессе эмбрионального развития. Подтверждением данного предположения служит

открытие в организме больных астмой аутоиммунных антител, специфически расщепляющих эндогенно образующийся пептидный гормон – вазоактивный кишечный пептид (VIP). По-видимому, по мере накопления экспериментальных данных число таких примеров будет увеличиваться. На скрытую потенциальную полифункциональность полипептидных цепей указывает также наличие у многих ферментов аллостерической регуляции, проявляющейся в различных формах, разная субстратная специфичность аллельных вариантов различных ферментов, а также изменение субстратной специфичности некоторых ферментов при изменении условий окружающей среды.

Таким образом, при ближайшем рассмотрении традиционное представление о ферментах как о высокоспецифических биологических катализаторах кажется упрощенным. Высокая специфичность ферментов проявляется лишь в определенных условиях внутренней среды организма, как одна из возможностей, детерминированных конкретной первичной структурой данного полипептида. Незначительные изменения первичной структуры таких полипептидных цепей под действием точковых мутаций, по-видимому, могли бы способствовать проявлению скрытых возможностей полипептидных цепей, вызывая появление у них новой субстратной специфичности. Реализация некоторых вновь возникших субстратных специфичностей мутантных ферментов может происходить в клетках в том случае, если для них найдется соответствующий субстрат среди многочисленных нормальных метаболитов и не менее многочисленных макромолекул (что имеет место, в частности, в случае аутоиммунных антител, расщепляющих VIP).

Таким образом, вырисовывается следующая гипотетическая цепь событий, приводящих к образованию ферментов с новой субстратной специфичностью: соматическая мутация изменяет первичную структуру полипептидной цепи какого-либо фермента таким образом, что он начинает распознавать эндогенный метаболит в качестве нового субстрата и метаболизировать его с образованием нового химического соединения, не свойственного данному организму. Эндогенно синтезированному химическому соединению можно было бы присвоить парадоксальное название *эндогенного ксенобиотика*. Естественно, что мутационное изменение субстратной специфичности ферментов должно быть редким событием. Однако с учетом громадного числа соматических клеток организмов животных и человека, а

также вполне значимой частоты спонтанных соматических мутаций (10^{-8} – 10^{-9} и выше на клеточную генерацию) этот механизм может реализовываться.

Возможность такого события должна существенно возрастать из-за увеличения частоты соматических мутаций под действием неблагоприятных экологических факторов. Более того, вышеупомянутые низкие частоты соматических мутаций можно рассматривать только как средние. В разделе 5.3 уже была обоснована дифференциальная защищенность отдельных генетических локусов от химических мутагенов, которая обуславливается особенностями их пространственной структуры в составе хроматина.

Если предположение об образовании эндогенных ксенобиотиков в онтогенезе верно, то подобный процесс должен быть свойственен большинству многоклеточных организмов и носить глобальный характер. В соответствии с этой гипотезой процесс образования эндогенных ксенобиотиков в многоклеточном организме ферментами, субстратная специфичность которых изменена под действием соматических мутаций, можно было бы назвать *ксенобиозом*. Какими же могут быть последствия ксенобиоза для многоклеточного организма? В качестве единственного примера рассмотрим возможное влияние соматических мутаций на систему цитохрома Р-450 в организме человека и животных.

Цитохромы группы Р-450 являются одними из основных компонентов обширной системы детоксикации чужеродных химических соединений (ксенобиотиков), поступающих в организм человека и животных из окружающей среды (подробнее см. раздел 5.1.1). Отличительной особенностью системы цитохромов Р-450 является их способность распознавать в качестве субстратов и метаболизировать множество экзогенных химических соединений, что, с одной стороны, объясняется наличием большого числа изоформ цитохромов Р-450, а с другой – их широкой субстратной специфичностью.

Среди цитохромов Р-450 встречаются изоформы и со строгой субстратной специфичностью. Эти ферменты участвуют в метаболизме эндогенных химических соединений, в частности в биосинтезе стероидных гормонов. Таким образом, налицо противоречивый характер действия цитохромов группы Р-450. С одной стороны, они распознают и метаболизируют миллионы ксенобиотиков, а с другой – как правило, не метаболизируют сотни тысяч эндогенных метаболитов. Следовательно, система цитохрома Р-450

толерантна в отношении эндогенных метаболитов собственного организма, многие из которых обладают большим структурным сходством с экзогенными ксенобиотиками. Каким же образом можно представить себе механизм потери толерантности системы цитохрома Р-450 по отношению к обычным метаболитам соматических клеток, и каковы возможные последствия этого явления? Метаболизм ксенобиотиков цитохромами группы Р-450 сопровождается двойственным физиологическим эффектом. Во-первых, как уже упоминалось (см. раздел 5.1.1), происходит детоксикация ксенобиотиков, а во-вторых, их метаболическая активация – образование нестабильных, реакционноспособных промежуточных химических соединений, обладающих мутагенной и канцерогенной активностью. Этот механизм является одним из основных в активации химических проканцерогенов в канцерогены, особенно в случае полициклических ароматических углеводородов.

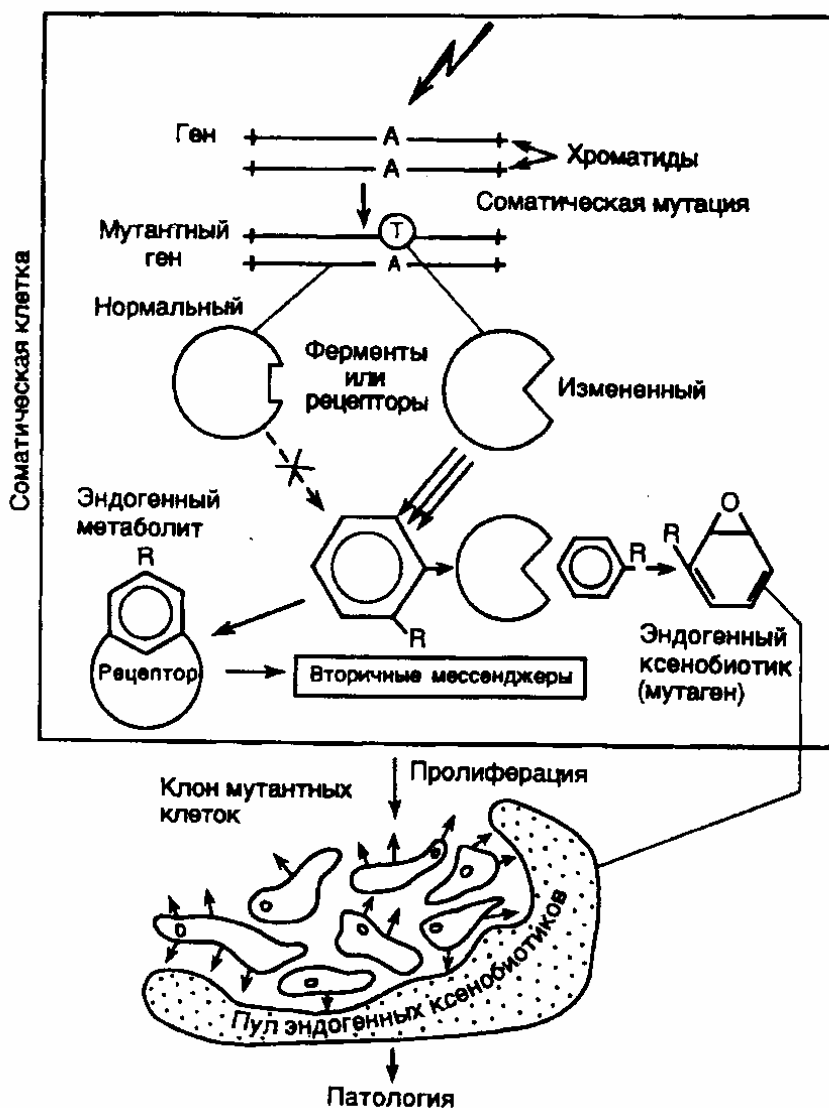


Рис. II.24. Гипотетическая схема ксенобиоза и его последствий

В верхней части рисунка изображены соматическая клетка и происходящие в ней процессы, приводящие к образованию эндогенных ксенобиотиков

При этом можно говорить о возможности нарушения толерантности системы цитохрома Р-450 в результате каких-либо патологических процессов, как это имеет место в случае иммунной системы при развитии аутоиммунных заболеваний. С учетом приведенных фактов можно предполагать, что под действием соматических мутаций может произойти изменение активного

центра цитохрома Р-450 таким образом, что он начнет распознавать в качестве субстрата эндогенный метаболит организма, например ароматическую аминокислоту или стероидный гормон, и метаболизировать его по эпоксид-диольному пути с образованием промежуточного эпоксида, обладающего мутагенной активностью (рис. II.24). Образовавшийся эндогенный ксенобиотик в том случае, если он не окажется особенно токсичным, будет непрерывно мутагенизировать геном как своей собственной клетки, так и соседних клеток до тех пор, пока не произойдут мутации, нарушающие контроль клеточной пролиферации (например в онкогенах или генах-супрессорах опухолей), что будет сопровождаться малигнизацией клеток, в том числе их неконтролируемым ростом, т.е. развитием злокачественного новообразования.

В том случае, если будет малигнизироваться клетка, содержащая мутантный ген цитохрома Р-450, то все клетки образующейся опухоли будут синтезировать эндогенный ксенобиотик и опухоль будет оказывать сильный токсический эффект на целый организм. Этого не произойдет, если будут малигнизироваться клетки, расположенные по соседству с мутантной. В настоящее время уже описаны эксперименты, в которых с помощью направленного введения точковых мутаций удалось изменить специфичность цитохрома Р-450. Кроме того, как было упомянуто выше, изменение субстратной специфичности под действием точковых мутаций описано и для других ферментов. Все это указывает на возможность развития событий по предлагаемому гипотетическому сценарию.

Появление мутаторного фенотипа у клеток злокачественных новообразований описано в настоящее время для многих форм рака. Основной причиной мутаторного фенотипа раковых клеток считают мутационные нарушения ферментных систем репарации, что имитирует внутриклеточное возрастание концентрации мутагенов. Предлагаемый механизм указывает на еще один возможный путь малигнизации нормальных клеток через образование эндогенных ксенобиотиков – эндогенных мутагенов. При развитии такой цепи событий повреждение систем репарации будет вторичным по отношению к первичному синтезу эндогенного мутагена.

Помимо онкологических заболеваний эндогенные ксенобиотики могли бы вызывать некоторые аутоиммунные заболевания в ответ на образование аддуктов с макромолекулами организма, что придавало бы модифицированным

макромолекулам иммуногенность и нарушало иммунологическую толерантность организма. Эндогенные ксенобиотики могли бы быть также одной из постоянно действующих причин старения организма вследствие его постепенного отравления эндогенными ксенобиотиками по мере накопления в клетках с возрастом мутантных ферментов, обладающих измененной субстратной специфичностью.

Процесс спонтанного и индуцированного соматического мутагенеза происходит непрерывно на протяжении всей жизни многоклеточного организма. Следовательно, в организме постоянно образуются новые мутантные белки, которые исходно не были закодированы его геномом. Мутантные белки могут контактировать в клетках с большим количеством эндогенных метаболитов и при наличии структурного соответствия между метаболитом и новым активным центром мутантного белка метаболизировать его с образованием эндогенного ксенобиотика, например продуктов расщепления упомянутого выше пептидного гормона VIP, наличие которых в организме больных доказано. Многоклеточный организм является своеобразным полигоном, на котором ежеминутно на протяжении всей жизни происходят испытания вновь образующихся белков на способность метаболизировать эндогенные химические соединения, с которыми они находятся в контакте. Таким образом, ксенобиоз в том случае, если он распространен так широко, как я думаю, должен быть обычным феноменом, присущим всем многоклеточным организмам.

Если предположение о ксенобиозе верно, то организм должен обладать эффективными системами защиты от этого неизбежного явления. Помимо высокоэффективных репаративных систем, понижающих частоту спонтанного мутагенеза в соматических клетках, защиту от эндогенных ксенобиотиков отчасти могла бы осуществлять сама иммунная система. В таких случаях с участием иммунной системы элиминировались бы те мутантные соматические клетки, которые экспонируют мутантные белки на своей поверхности и которые в этом случае распознавались бы как чужеродные антигены. Однако в большинстве случаев, по-видимому, этого не должно происходить, так как мутантные белки будут локализованы внутри клеток.

Другим средством защиты от образующихся ферментов с новой субстратной специфичностью (и ксенобиоза) у многоклеточных организмов могут быть убиквитин- и АТР-зависимые системы внутриклеточного протеолиза

аномальных белков. Однако эти системы должны срабатывать в первую очередь в отношении мутантных белков, у которых резко нарушена пространственная структура, что при одиночных заменах аминокислот, по-видимому, происходит редко. В пользу такого предположения свидетельствуют все те многочисленные мутантные соматические клетки, которые описаны в литературе и из которых можно выделять мутантные белки.

Вероятно, основную защиту от эндогенных ксенобиотиков могла бы обеспечивать обсуждавшаяся выше система цитохрома Р-450. Действительно, широкая совокупная субстратная специфичность системы цитохрома Р-450 позволяет ей метаболизировать и в конечном счете обезвреживать химические соединения. По аналогии с иммунной системой, которая помимо защиты организма от экзогенных инфекционных агентов выполняет функции иммунного надзора за антигенным состоянием соматических клеток, основной функцией системы цитохрома Р-450 может быть освобождение организма от эндогенных ксенобиотиков. В соответствии с этим система цитохрома Р-450 эволюционно могла возникнуть для поддержания, прежде всего, метаболического гомеостаза многоклеточных организмов в условиях непрерывно образующихся эндогенных ксенобиотиков, а уже затем ее функции распространились и на экзогенные химические соединения, вредные для организма.

ГЛАВА 9. АНТИСМЫСЛОВЫЕ РНК, РИБОЗИМЫ И ДЕЗОКСИРИБОЗИМЫ

Антисмысловые РНК как ключевые компоненты одной из систем негативной регуляции экспрессии генов были впервые описаны у бактерий. Вскоре тот же тип регуляции был обнаружен и у эукариот. Уже в первых опытах было установлено, что короткие РНК, комплементарные мРНК, образуют с ними гибриды и блокируют трансляцию. Поскольку действие таких РНК направлено против функционирования кодирующих (осмысленных) РНК, они получили название *антисмысловых* (antisense RNA), или micРНК (mRNA interfering complementary RNA). Это открытие вскоре легло в основу целого направления исследований искусственной регуляции экспрессии генов, неразрывно связанного с методами генной инженерии.

9.1. Антисмысловые РНК и олигонуклеотиды

Главный механизм, лежащий в основе функционирования системы антисмысловых РНК, прост и опирается на известный феномен взаимодействия двух комплементарных друг другу молекул нуклеиновых кислот с образованием двухцепочечных РНК–РНК- или ДНК–РНК-гибридов. Оказалось, что взаимодействие с мРНК комплементарного ей полинуклеотида или олигонуклеотида (которые могут быть транскриптами незначущей цепи ДНК, т.е. противоположной той, с которой произошла транскрипция этой мРНК) может блокировать ее трансляцию рибосомами и нарушать экспрессию всего гена на уровне трансляции.

В середине 1970-х годов синтетические олигонуклеотиды, комплементарные мРНК, были впервые использованы в бесклеточных системах биосинтеза белка для подавления трансляции этих мРНК. С разработкой современных методов генной инженерии получение антисмысловых РНК упростилось, так как достаточно в экспрессирующем векторе поместить кодирующую последовательность гена в обратной ориентации по отношению к промотору, чтобы начала транскрибироваться незначущая цепь ДНК с образованием антисмысловой РНК. В такой системе с использованием современных векторов, в которых клонируемая

последовательность фланкирована промоторами T3-, T7- или SP6-РНК-полимераз (например вектор Bluescript, см. рис. II.5), можно синтезировать большое количество антисмысловых РНК и использовать их в высокоочищенном состоянии в опытах *in vitro*. Однако современные генно-инженерные конструкции позволяют вводить векторные молекулы, экспрессирующие антисмысловые РНК, непосредственно в клетки живых организмов и наблюдать биологические эффекты антисмысловых РНК в результате их эндогенной экспрессии.

9.1.1. Механизм действия антисмысловых РНК

Многочисленные исследования антисмысловых РНК как *in vitro*, так и *in vivo* показали, что конечным результатом их действия, как правило, является высокоспецифическое ослабление экспрессии генов, мРНК которых является мишенью действия антисмысловых РНК. При использовании олигодезоксирибонуклеотидов, комплементарных различным частям мРНК, было установлено, что как в бесклеточных экстрактах, так и в эукариотических клетках подавление трансляции мРНК достигалось главным образом за счет ее расщепления РНКазой H в месте образования РНК–ДНК-гибрида. При этом бесклеточные экстракты зародышей пшеницы были особенно богаты РНКазой H, однако процесс специфического расщепления мРНК эффективно происходил и в ооцитах *Xenopus*, и в культивируемых клетках животных. Интересным следствием такого рода исследований оказалось то, что даже экзогенно добавленные к культивируемым клеткам олигодезоксирибонуклеотиды проникали внутрь клеток и блокировали трансляцию соответствующих мРНК.

Механизм транспорта олигонуклеотидов в клетки обладает чертами сходства с эндоцитозом, опосредованным рецепторами. Эффект подавления трансляции мРНК специфическими комплементарными нуклеотидами достигался, как правило, лишь при большом молярном избытке олигонуклеотидов по отношению к мРНК, что является следствием их относительно медленной реассоциации и быстрой внутриклеточной деградации олигонуклеотидов (время полужизни в ооцитах менее 20 мин). Расщепление гибридов мРНК с олигодезоксирибонуклеотидами – не единственный механизм ингибирующего действия олигонуклеотидов на трансляцию. Использование

метилфосфонатных ($-P-CH_3$) аналогов олигонуклеотидов, гибриды которых с РНК не расщепляются РНКазой Н, показало, что их действие может быть следствием изменения пространственной структуры мРНК, а также нарушения ее связывания рибосомами. При этом олигонуклеотиды, в которых чередуются фосфодизэфирные и метилфосфонатные связи, сохраняли способность индуцировать расщепление мРНК РНКазой Н в гибридах и в то же время характеризовались повышенной устойчивостью к действию внутриклеточных нуклеаз.

Использование аналогов олигодезоксирибонуклеотидов повышает результативность действия системы ингибирования трансляции. Еще более эффективными являются системы, в которых определенный уровень антисмысловых РНК непрерывно поддерживается за счет их эндогенного синтеза, а с мРНК взаимодействуют более протяженные комплементарные последовательности нуклеотидов.

Механизмы ингибирующего действия антисмысловых РНК. В первых опытах с антисмысловыми РНК, которые с помощью микроинъекции вводили в ооциты *X. laevis*, было установлено, что они подавляют трансляцию соответствующих мРНК. В том случае, если антисмысловые РНК, комплементарные мРНК креатинкиназы В человека, синтезировались в ядрах клеток лимфомы человека, сплайсинг, 3'-концевой процессинг и экспорт соответствующих мРНК в цитоплазму не были нарушены, хотя антисмысловые РНК синтезировалась в эквимольных количествах по отношению к мРНК-мишени, а сама антисмысловая РНК была комплементарна последовательностям нуклеотидов части последнего кодирующего экзона, соседнего интрона и 3'-фланкирующей последовательности. Экспорт антисмысловой РНК из ядра в цитоплазму был несколько снижен по сравнению с соответствующей мРНК, хотя при этом происходило уменьшение внутриклеточной активности креатинкиназы на 40%. Образование дуплекса между антисмысловой РНК и мРНК, по-видимому, сопровождалось ингибированием связывания мРНК рибосомами, а также транслокации рибосом в процессе трансляции. Взаимодействие антисмысловых РНК с 5'-концевыми нетранслируемыми последовательностями мРНК может изменять пространственную структуру мРНК таким образом, что иницирующий AUG-кодон становится недоступным рибосомам. Не исключено, что в ряде случаев

антисмысловые РНК могут блокировать экспорт мРНК из ядра в цитоплазму, особенно когда они присутствуют в молярном избытке. Это также может быть одной из причин подавления синтеза соответствующих белков.

Наилучшие результаты по блокированию экспрессии генов были получены в том случае, когда гены антисмысловых РНК находились в клетках в виде множественных экспрессируемых копий, встроенных в геном, или под контролем сильных промоторов. Дуплексы антисмысловых РНК и мРНК, образующиеся в клетках, которые экспрессируют антисмысловые РНК, являются субстратом для расплетающего фермента (*unwindase*), который присутствует в клетках всех типов тканей животных. Расплетающий фермент производит дезаминирование 25–40% остатков аденина в обеих цепях дуплекса с образованием инозина. Остатки инозина образуют комплементарные пары с остатками гуанина, поэтому в процессе такой модификации происходит изменение кодирующих свойств мРНК и ее трансляция может приводить к образованию нефункциональных белков. Таким образом, и в данном случае антисмысловые РНК опосредуют инактивацию комплементарных им мРНК, что является еще одним механизмом их ингибирующего действия на экспрессию генов.

9.1.2. Использование антисмысловых РНК

Получение фенокопий. Клетки или организмы, обладающие фенотипом мутантных клеток или организмов, сформировавшимся не вследствие мутаций, называют *фенокопиями*. Развитие техники антисмысловых РНК позволило исследовать функции отдельных клонированных генов, дифференциально экспрессирующихся в процессе онтогенеза животных. Экспрессия антисмысловых РНК в клетках организма на определенных стадиях его развития может сопровождаться понижением внутриклеточного уровня соответствующих белков или ферментов, что имитирует процесс мутационной инактивации их генов. Одним из примеров использования антисмысловых РНК для получения фенокопий у мышей были исследования роли гена основного белка миелина (ОБМ) в онтогенезе.

Получение трансгенных мышей, экспрессирующих антисмысловые РНК гена ОБМ, который содержался в количестве десяти копий на геном, сопровождалось понижением на 80% внутриклеточного содержания этого

белка. Примерно половина потомства мышей приобретала фенотип *shiverer* через две недели после рождения. Такой фенотип, для которого характерна непрерывная дрожь конечностей, был впервые описан у мышей, гомозиготных по рецессивной мутации в гене ОБМ. Однако для мутантных фенотипов подобных фенотипов была характерна значительная вариабельность, а также имели место мозаицизм в распределении в мозге клеток, содержащих ОБМ (т.е. лишь часть клеток содержала ОБМ), и гетерогенность в отношении уровня экспрессии этого гена. В табл. II.3 суммированы результаты экспериментов по получению фенотипов у мышей, проведенных к 1996 г.

Антисмысловые РНК могут быть использованы и для получения мутантных фенотипов, не описанных ранее из-за отсутствия соответствующих генетических мутантов. Например, введение посредством трансгеноза самкам дрозофилы антисмыслового гена рибосомного белка RpA1, находящегося под контролем промотора теплового шока, сопровождалось нарушением оогенеза при повышенной температуре (37°). При 18°, когда экспрессия антисмысловых РНК резко снижена, уменьшалось и образование дефектных яиц. Проявление мутантного фенотипа в этом случае зависело от времени индукции антисмысловых РНК, дозы (количества) антисмыслового гена и уровня его транскрипции. Таким образом, в данной работе было подтверждено известное наблюдение, что для проявления мутантного фенотипа в некоторых случаях нет необходимости в полной инактивации функции соответствующего гена.

Таблица II.3

**Влияние экспрессии антисмысловых РНК на фенотип
трансгенных мышей**

Гены-мишени	Длина miсРНК, п.о.	Мишень в гене	Фенотип
Основной белок миелина	1200	Экзоны	Понижение уровня мРНК на ≤80%; мыши <i>shiverer</i>
ГФРТ	1390	5'-НТ, экзон 1 и интрон 1	Понижение уровня мРНК на 20–50%; активность фермента не изменилась
»	550	То же	Эффект отсутствовал
ψ-Последовательность мышиного вируса лейкемии Молони	540	ψ-Область	Снижение числа лейкозов у зараженных мышей с 31% до 0%
Рецептор глюкокортикоидов типа II	1815	3'-НТ	Понижение уровня мРНК в мозге на 50–70%; уменьшение содержания рецептора
β-Цепь иммуноглобулина А	100	5'-НТ и инициирую- щий кодон	Понижение уровня мРНК на ≥50%; задержка созревания В- клеток
Фактор роста нервов	1300	Экзоны	Изменение базального уровня фактора роста нервов в коже
<i>wnt-1</i>	4700	Экзоны	Понижение уровня мРНК на ≤98%
Gα _{i2}	39	5'-НТ и инициирую- щий кодон	Понижение уровня белка в печени и жировой ткани на ≤95%, уменьшение скорости роста тела

Таблица II.3 (окончание)

Гены-мишени	Длина miсРНК (п.о.)	Мишень в гене	Фенотип
Белок нуклеокапсида вируса гепатита мышей	1800	Экзоны	Повышение устойчивости к вирусной инфекции
<i>GLUT-2</i>	1489	»	Понижение уровня белка <i>GLUT-2</i> на 80%
Интерлейкин 3	990	»	Понижение уровня IL-3, приводящее к развитию лимфопролиферативных заболеваний β -клеток
Миниген COLIA1	~1200	»	Понижение уровня белка COLIA1 на 50%, приводящее к летальной ломкости костей
Ангиотензиноген	1900	»	До индукции промотора – повышение уровня мРНК в печени; после индукции – понижение уровней мРНК в печени и почках и повышение в мозге на 20 дней, после чего уровни нормализуются; понижение уровня ангиотензиногена в плазме через 30 дней после индукции
Глюкокиназа β -клеток	1900	»	Понижение уровня мРНК в β - клетках, сопровождаемое 50%- ным снижением активности глюкокиназы, повышение уровня глюкозы в крови

Использование антисмысловых РНК для исследования онтогенеза животных выявило высокую чувствительность определенных его стадий к изменению уровня экспрессии генов, а также неспецифическую токсичность антисмысловых транскриптов, проявляющуюся за счет взаимного влияния различных частей ранних эмбрионов друг на друга и гетерогенности экспрессии самих антисмысловых РНК в развивающихся зародышах. Это затрудняет интерпретацию результатов, полученных при экспрессии антисмысловых РНК в раннем эмбриогенезе. Поэтому более удобными объектами для получения фенокопий являются культивируемые соматические клетки животных и растений.

Исследование клеточного цикла. Антисмысловые РНК оказались полезными в исследованиях роли протоонкогенов и факторов роста в пролиферации клеток. С использованием такого подхода были изучены протоонкогены *c-fos*, *c-myc*, *c-src*, антионкоген *p53*. В норме уровень экспрессии *c-myc* выше в пролиферирующих клетках, чем в дифференцирующихся. В присутствии антисмысловых олигонуклеотидов наблюдали понижение внутриклеточного уровня белка с-Мус в клетках линии HL60, что сопровождалось понижением скорости роста клеток и усилением миелоидной дифференцировки. Природные антисмысловые с-мус-РНК были обнаружены среди транскриптов клеток мышей. Антисмысловые *c-fos*-РНК предотвращали повторное вхождение покоящихся клеток 3Т3 (фаза G₀) в клеточный цикл.

Противовирусная терапия. Внутриклеточная экспрессия антисмысловых РНК, направленных против мРНК некоторых ключевых вирусных белков, может эффективно подавлять вирусную инфекцию. Этот результат был впервые получен с клетками *E. coli*, содержащими плазмиды, которые экспрессировали антисмысловые РНК, комплементарные различным участкам генов бактериофага SP. Бактериальные клетки приобретали иммунитет к фагу в наибольшей степени в том случае, если антисмысловые РНК были комплементарны 5'-концевым последовательностям мРНК (включая последовательности Шайна–Дальгарно) гена белка созревания. Менее эффективными оказались антисмысловые РНК, комплементарные мРНК белков оболочки бактериофага или его репликазы. 15-Нуклеотидная антисмысловая последовательность, комплементарная сайту связывания рибосом мРНК, на 94% подавляла выход фага SP и оказывала значительный ингибирующий

эффект на развитие родственных фагов Q β и GA. Эти опыты показали принципиальную возможность использования антисмысловых РНК для борьбы с вирусными инфекциями. Кроме того, высокая специфичность антисмысловых РНК по отношению к вирусным последовательностям нуклеотидов сводит к минимуму возможные цитотоксические и другие побочные действия. Современные методы генной инженерии позволяют вводить гены антисмысловых РНК в клетки зародышевой линии животных и растений, что может обеспечивать наследование этих генов в ряду поколений.

Обнадёживающие результаты были получены в опытах по ингибированию развития вируса саркомы Рауса (BCP) в клетках перепелов. Антисмысловые РНК, комплементарные мРНК белка оболочки BCP, на 70–80% предотвращали внутриклеточное развитие BCP, дефектного по белку оболочки, в присутствии плазмиды, экспрессирующей этот белок, которая без антисмысловой РНК комплементировала развитие дефектного вируса.

Были получены трансгенные растения табака, которые содержали экспрессирующий вектор, направляющий синтез антисмысловой РНК, комплементарной мРНК белка оболочки вируса мозаики огурцов (ВМО). Такие растения оказались устойчивыми к ВМО только при низкой множественности инфекции, а иммунитет к ВМО мог быть преодолен с помощью высоких концентраций инокулируемого вируса. Аналогичные результаты были получены с трансгенными растениями табака, экспрессирующими антисмысловые РНК, комплементарные транскрипту гена белка оболочки X. В этих опытах уровень экспрессии антисмысловых РНК оказался недостаточным для того, чтобы обеспечить полную устойчивость растений к вирусной инфекции, что требует дальнейшей оптимизации условий экспрессии генов антисмысловых РНК в клетках растений.

9.1.3. Природные антисмысловые РНК

За то время, которое прошло с момента открытия в середине 1970-х годов антисмысловых РНК и их успешного использования для искусственной регуляции экспрессии генов, стало ясно, что этот эффективный генно-инженерный прием уже давно и широко используется самой природой для тех же целей – модуляции экспрессии генов на молекулярном уровне. В

прокариотических системах антисмысловые РНК участвуют в регуляции репликации и поддержании плазмид. Так, в случае плазмиды ColEI инициация ее репликации негативно регулируется с помощью короткой, нетранслируемой антисмысловой РНК. Антисмысловая РНК I плазмиды ColEI также играет ключевую роль в обеспечении несовместимости плазмид, принадлежащих к одной группе совместимости, внутри одной бактериальной клетки (подробнее см. раздел 4.2.2). Похожие механизмы с участием антисмысловых РНК используются и для регуляции репликации плазмид группы IncF1 (NR1, R1 и R6) и IncQ (R1162), а также для регулирования числа копий плазмиды pT181 *Staphylococcus aureus*. Антисмысловые РНК служат у бактерий для регуляции экспрессии генов на уровне транскрипции (например гена белка-рецептора циклического AMP) и трансляции. В последнем случае антисмысловая РНК *micF* участвует в подавлении экспрессии белка *OmpF* внешней мембраны *E. coli*.

Экспрессия гена *sulA* *E. coli*, кодирующего SOS-индуцируемый ингибитор деления бактериальных клеток, регулируется нетранслируемой РНК, комплементарной на протяжении 250 нуклеотидов 3'-концевой части *sulA*-РНК. Регуляция на уровне трансляции антисмысловыми РНК имеет место у колицинового (*E1*) оперона плазмиды ColEI, а также оперона *gvpABC* цианобактерии *Calotrix 7601*, которые обеспечивают образование газовых везикул, придающих плавучесть бактериальным клеткам.

Антисмысловые РНК участвуют и в регуляции жизненного цикла бактериофагов. Например, у бактериофага λ обнаружена короткая антисмысловая РНК, комплементарная мРНК гена *Q* и ингибирующая ее трансляцию. Это приводит к понижению уровня экспрессии поздних генов бактериофага и ингибированию его литического развития. Другой антисмысловой РНК фага λ , участвующей в регуляции экспрессии гена *cII*-репрессора, является оорРНК. Эта РНК комплементарна 3'-концевой части транскрипта гена *cII* и после образования гибрида ингибирует ее трансляцию, вероятно, вследствие расщепления дуплекса РНКазой III клетки-хозяина. Небольшая антисмысловая РНК (*sar*РНК) участвует в регуляции синтеза антирепрессора фага P22. Та же система регуляции антисмысловых РНК характерна и для некоторых транспозонов. Антисмысловые РНК природного происхождения были обнаружены и в клетках эукариот. Так, наличие

небольших ядерных поли(A⁻)-РНК, комплементарных гену дигидрофолатредуктазы, характерно для ядер клеток мышей. Геномные локусы с генами, транскрибируемыми в двух противоположных направлениях, имеются у мышей, крыс и дрозофилы. Образование антисмысловых РНК отмечено во время латентной инфекции вирусом простого герпеса, а также в семенах ячменя. В последнем случае были идентифицированы антисмысловые РНК, комплементарные мРНК изозимов α -амилазы типов А и В.

Антисмысловые РНК, по-видимому, играют важную роль и в экспрессии ранних генов вируса полиомы. Геном вируса полиомы представляет собой небольшую кольцевую двухцепочечную молекулу ДНК, на которой транскрипция ранних и поздних генов происходит в противоположных направлениях с промоторов, расположенных в межгенной регуляторной области. Соответственно, в результате экспрессии ранних и поздних генов транскрибируются разные цепи ДНК. Терминация транскрипции поздних генов происходит неэффективно на поздних стадиях инфекции, что приводит к образованию в ядрах зараженных клеток гигантских РНК, заключающих в себе многократно повторяющуюся последовательность всего генома. При этом последовательности нуклеотидов таких РНК, соответствующие ранним генам вируса, являются антисмысловыми по отношению к нормальным ранним РНК вируса как продуктам транскрипции противоположной цепи вирусной ДНК. Накопление таких антисмысловых РНК приводит к резкому снижению внутриклеточного уровня ранних РНК, а экспериментальная дестабилизация антисмысловых РНК с помощью мутаций сопровождается значительным его повышением *in vivo*.

Приведенных примеров достаточно, чтобы сделать вывод о широком распространении антисмысловых РНК в природных условиях, хотя в случае эукариотических организмов физиологическое значение антисмысловых РНК непонятно. Расширение исследований в этой области молекулярной генетики несомненно приведет к открытию новых регуляторных механизмов с участием антисмысловых РНК. В том случае, если регуляторные функции антисмысловых РНК, реализуемые через РНК-РНК-гибриды, будут подтверждены у эукариот, это еще раз укажет на перспективность направления поисков путей искусственной регуляции экспрессии эукариотических генов с

использованием антисмысловых РНК у животных и растений. Правильность такого направления уже сейчас подтверждается возможностью получения фенокопий у высших организмов с помощью антисмысловых РНК, а для повышения эффективности их действия необходимо просто оптимизировать условия их биосинтеза в эукариотических клетках. Использование техники антисмысловых РНК позволяет осуществлять высокоспецифическую инактивацию функционирования определенных генов *in vivo* и становится особенно полезным в тех случаях, когда инактивация соответствующих генов генетическими методами (с помощью мутаций) невозможна. Тем не менее значительная вариабельность в уровнях экспрессии антисмысловых РНК в клетках, содержащих эндогенные антисмысловые гены, а также некаталитический характер действия антисмысловых РНК, требующий их высокой внутриклеточной концентрации, накладывают серьезные ограничения на их применение. В этой связи перспективным направлением представляется использование рибозимов, фланкированных последовательностями антисмысловых РНК. Внутриклеточная стабильность таких макромолекул может быть существенно повышена путем конструирования антисмысловых РНК с оптимально подобранной пространственной структурой, а также использованием аналогов нуклеотидов для модификации этих макромолекул.

9.1.4. Антисмысловые РНК и патология: возможный механизм возникновения доминантных мутаций

Вся совокупность полученных к настоящему времени данных позволяет предполагать, что антисмысловые РНК могут образовываться не только в результате реализации нормальных регуляторных механизмов, но и во время различных патологических состояний организма, выступая одной из причин их возникновения. Действительно, для того чтобы в клетке начали образовываться антисмысловые РНК, достаточно возникновения инверсии кодирующей части экспрессируемого гена, расположенной за промотором, направляющим транскрипцию этого гена. В таком случае может осуществляться инактивация самого гена, в котором произошла инверсия, и будет образовываться антисмысловая РНК, комплементарная нормальной мРНК мутировавшего гена. К аналогичным результатам будут приводить и транслокации кодирующих частей генов (с одновременной инверсией) под контроль промоторов других

активно экспрессирующихся генов. По своей природе такого рода мутации, приводящие к образованию антисмысловой РНК, должны сопровождаться подавлением экспрессии аллельного гена, расположенного на гомологичной хромосоме, и даже целого семейства гомологичных генов. Следствием этого должно быть резкое снижение уровня и даже полное отсутствие в мутантной клетке белкового продукта мутантного гена с образованием фенокopies. К каким же последствиям для клетки могут приводить такие мутации?

В качестве одного из примеров, демонстрирующих потенциальную важность этих мутаций в развитии патологических состояний организма, рассмотрим возможные последствия инактивации с помощью вышеупомянутого механизма генов-супрессоров опухолевого роста, называемых также *рецессивными онкогенами* или *антионкогенами*. В настоящее время можно считать доказанным, что различные формы такого онкологического заболевания, как ретинобластома, возникают вследствие инактивации в единственной клетке-предшественнице обеих копий одного гена – *RB1*. В случае наследуемых форм ретинобластомы один мутантный аллель гена *RB1* наследуется от одного из родителей больного, тогда как другой образуется в результате соматической мутации в малигнизирующейся клетке-предшественнице опухоли. При спорадической ретинобластоме инактивация обоих аллелей гена должна произойти в одной клетке, что является редким событием. Предлагаемая модель с участием антисмысловых РНК объясняет, каким образом инактивация обоих аллельных генов *RB1* может происходить в результате одного мутационного события: инверсии части кодирующей последовательности гена *RB1* или одновременной инверсии и транслокации части гена в новый хромосомный локус под контроль сильного промотора другого гена. В настоящее время у млекопитающих известны не менее 12 хромосомных локусов, утрата которых в результате делеций сопровождается малигнизацией клеток и которые содержат другие гены-супрессоры опухолей. Каждый из этих локусов потенциально может быть инактивирован с помощью эндогенных антисмысловых РНК. Следовательно, концепция антисмысловых РНК дает простое объяснение молекулярных механизмов одновременной инактивации аллельных рецессивных онкогенов в живом организме с последующей малигнизацией клеток и развитием опухолей.

9.2. Рибозимы и дезоксирибозимы

Регуляция экспрессии генов с помощью антисмысловых РНК характеризуется высокой специфичностью. Это обусловлено большой точностью процесса РНК-РНК-гибридизации, основанной на комплементарном взаимодействии друг с другом протяженных последовательностей нуклеотидов. Однако сами по себе антисмысловые РНК не инактивируют необратимо мРНК-мишени, и для подавления экспрессии соответствующих генов требуются высокие (по крайней мере, эквимоларные по отношению к мРНК) внутриклеточные концентрации антисмысловых РНК. Эффективность действия антисмысловых РНК резко повысилась после того, как в их состав были введены молекулы *рибозимов* – коротких последовательностей РНК, обладающих эндонуклеазной активностью. Хотя ниже термин "рибозимы" будет использоваться почти исключительно в отношении молекул РНК, обладающих эндонуклеазной активностью, необходимо иметь в виду, что в настоящее время известно множество других ферментативных активностей, ассоциированных с РНК. Поэтому рибозимами в широком смысле называют молекулы РНК, обладающие *любой* ферментативной активностью. Некоторые из этих активностей будут рассмотрены в разделе 9.3 при обсуждении возможной роли РНК в происхождении жизни.

9.2.1. Типы рибозимов

Эндорибонуклеазная активность РНК была впервые обнаружена Т. Чехом в 1980 г. у интрона группы I предшественника рибосомной РНК Tetrahymena, осуществляющего аутокаталитическую реакцию сплайсинга (аутосплайсинга), в результате которой происходит вырезание из молекулы предшественника рРНК последовательности этого интрона и образование зрелой молекулы рРНК. (Молекулярный механизм реакции аутосплайсинга уже был рассмотрен в первой части книги в разделе 3.3.2.) С тех пор аутокаталитические реакции расщепления были выявлены у многих молекул РНК (рис. II.25). Наиболее перспективными для использования в составе антисмысловых РНК оказались рибозимы, образующие пространственную структуру типа "головки молотка" (hammerhead – HH) (см. рис. II.25,а). Такие РНК были найдены у сателлитных РНК вирусов растений, вироидов, а также

среди транскриптов сателлитных ДНК тритонов. У РНК подобного типа была определена минимальная каноническая последовательность, необходимая для появления ферментативной активности, длиной в 27 нуклеотидов, из которых высококонсервативны 13 нуклеотидов, а также последовательность из трех нуклеотидов, в пределах которой происходит расщепление фосфодиэфирной связи (см. рис. II.25,а).

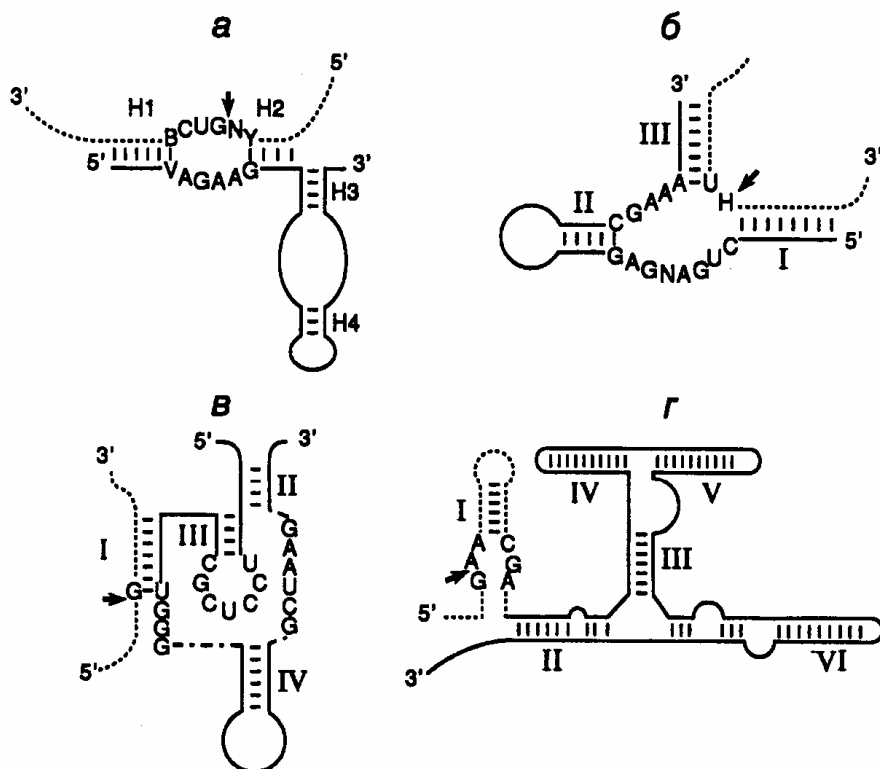


Рис. II.25. Особенности вторичной структуры рибозимов разных типов

а – "головка молотка", б – шпилька, в – рибозим вируса гепатита δ , г – рибозим *Neurospora* VS. Стрелки обозначают точки расщепления РНК; нуклеозид N может быть A, U, G или C, H – A, U или C, Y – любой пиримидин; приведены общепринятые обозначения элементов вторичных структур рибозимов

Другими примерами рибозимов являются РНК, образующие структуры типа шпилек, а также РНК вируса гепатита δ (ВГД) и сателлитная РНК Варкуда (*Varcud satellite* – VS) нейроспоры (см. рис. II.25,б,в,г соответственно). (Структура и механизм действия рибозимов в составе интронов групп I, II и III

были подобно рассмотрены в разделе 2.2.3.) Рибозимы со структурами типа шпилек были впервые обнаружены у виридов и вирусидов растений, рибозимы ВГД были идентифицированы среди сателлитных РНК вируса гепатита В человека, а рибозимы VS – среди сателлитных РНК природных изолятов *Neurospora*. В связи с рибозимами следует вспомнить и о РНКазе Р *E. coli*, состоящей из небольшого полипептида и РНК длиной в 375 нуклеотидов, о которой уже упоминалось в разделе 2.2.1. Именно РНК обеспечивает этому ферменту эндонуклеазную активность, необходимую для процессинга предшественников тРНК.

Хотя у природных РНК сайт расщепления и последовательность нуклеотидов рибозима расположены на одной молекуле (*in cis* по отношению друг к другу), НН-рибозимы могут действовать и на сайты расщепления, расположенные *in trans*, т.е. на других молекулах РНК. Именно это свойство рибозимов данного класса позволило использовать их для создания высокоспецифичных эндорибонуклеаз на основе антисмысловых РНК, которые обладают способностью расщеплять *in trans* молекулы РНК, содержащие тринуклеотидные сайты расщепления. Для создания высокоспецифического рибозима, функционирующего по этому принципу, последовательность рибозима заключают в последовательность антисмысловой РНК таким образом, чтобы при образовании гибрида с РНК-мишенью в контакте с активным центром рибозима оказывалась тринуклеотидная последовательность сайта расщепления.

9.2.2. Свойства рибозимов

Стабильность рибозимов в биологических жидкостях.

Нестабильность РНК является одним из основных ограничений, препятствующих эффективному их использованию *in vivo* в качестве лекарственных средств. Поскольку рибозимы представляют собой короткие молекулы, которые можно получать в результате химического синтеза, химические модификации этих олигонуклеотидов рассматриваются как одна из перспективных возможностей повышения их устойчивости к нуклеазам. Было установлено, что введение 2'-О-Ме-модифицированных нуклеотидов вместо обычных во все, кроме пяти, положения НН-рибозима позволило повысить его стабильность, по крайней мере, в 1000 раз без существенной потери

каталитической активности. Другие модификации нуклеотидов в определенных положениях: введение 2'-О-аллильных групп, фосфоротиоатных связей вместо фосфодиэфирных, замена пиримидинов их 2'-амино- или 2'-фторпроизводными, использование различных 2'-производных сахаров в сахарофосфатном остове РНК и т.п., как правило, приводили к снижению активности рибозимов, но сопровождались повышением их стабильности. Улучшенные конструкции рибозимов с модифицированными сахарами обладали каталитической активностью, свойственной рибозима́м дикого типа, со временем полужизни в сыворотке крови человека 5–8 ч. Кроме того, введение на 3'-конец рибозима остатка dT, соединенного связью 3'-3', сопровождалось увеличением времени его полужизни в сыворотке до 3 дней.

Адресная доставка искусственных рибозимов. Существуют два основных подхода к адресной доставке олигонуклеотидов. При одном из них синтезированные молекулы рибозимов объединяют химическими методами или физически с макромолекулами, например липидами или лигандами рецепторов. В зависимости от природы молекул, находящихся в комплексе с рибозимом, его доставка будет либо неспецифической (в составе липосом), либо более специфической, которая может обеспечиваться, например эндоцитозом, опосредованным рецепторами. Другая группа методов использует экспрессирующие векторы, содержащие ген рибозима, транскрипция которого сопровождается внутриклеточным биосинтезом его молекул.

В ходе исследований по оптимизации адресной доставки рибозимов были изучены многие конструкции липидных носителей. Такие конструкции позволяют инкапсулировать большое количество молекул рибозимов как в водном окружении, так и растворенными в липидном бислое. Соединение липосом с молекулами соответствующих антител или других лигандов может до некоторой (хотя и небольшой) степени обеспечивать тканеспецифическую доставку рибозимов и их проникновение в клетки требуемых типов, например пораженных вирусами. Циркулирующие липосомы быстро удаляются из кровотока макрофагами ретикулоэндотелиальной системы. Химическая модификация липидного бислоя или покрытие липосом оболочкой из полиэтиленгликоля ослабляют неспецифическую сорбцию белков сыворотки, а следовательно, и неспецифическое распознавание липосом макрофагами.

Для создания альтернативных гидрофобных комплексов молекул рибозимов с носителем иногда используют катионные липиды. Такие липиды содержат длинные цепи жирных кислот (обычно C_{16} или C_{18}), соединенные с полиамином. Полагают, что комплексы катионных липидов с олигонуклеотидами не образуют липосомоподобных структур. Результаты, полученные при исследовании возможностей адресной доставки рибозимов с помощью липосом и другими аналогичными методами, вселяют оптимизм. Однако до сих пор эти методы остаются чисто эмпирическими и должны разрабатываться индивидуально для каждого нового рибозима и типа клеток, в которые предполагается его доставлять.

Для адресной доставки генов рибозимов в составе экспрессирующих векторов используют те же приемы, что и при генотерапии (подробнее см. раздел 10.4). Особенно часто для этой цели применяются ретровирусы. Включение в состав рибозима сигнальной последовательности этих вирусов, обеспечивающей упаковку их РНК в вирусные частицы, на 90% снижает титр соответствующих ретровирусов. При этом рибозим не оказывал влияния на те же мишени, локализованные в цитоплазме.

Аденоассоциированные вирусы (ААВ), способные доставлять свою ДНК в клетки многих типов, также часто используются в качестве носителей для векторов, экспрессирующих рибозимы. Недавно сконструирован вектор, который из всех последовательностей ААВ содержал только концевые последовательности. С использованием этого вектора получена экспрессия антисмысловых РНК гена *tar* вируса иммунодефицита человека, которая блокировала его размножение. С помощью аденовирусов можно не только направленно осуществлять доставку экспрессирующихся генов, но и переносить большие молекулы декстранов, белков и плазмид, связанных с лигандами, способными и неспособными к независимой репликации. Например, удастся транспортировать гены к дыхательному эпителию в составе комплексов аденовирус–полилизин–ДНК. Адресная доставка векторов к требуемым клеткам или тканям может быть осуществлена также непосредственно в виде аэрозоля (эпителий легких) или *ex vivo* в случае клеток костного мозга, который далее повторно вводят в организм больного.

Исследование функционирования рибозимов *in vivo*. Для доказательства возможности использования рибозимов как лекарственных

средств необходимо было прежде всего продемонстрировать, что осуществляемое ими расщепление мРНК сопровождается ожидаемыми физиологическими и биохимическими изменениями в организме. В одних из первых опытов использовали рибозимы, направленные против мРНК хлорамфениколацетилтрансферазы (ХАТ), ген которой экспрессировался в клетках животных. Рибозимы, синтез которых обеспечивался экспрессирующими векторами, специфически на 60% подавляли внутриклеточный синтез ХАТ. Кроме того, было показано направленное действие рибозимов против 28S рРНК и ее предшественников, а также мРНК α -лактальбумина, но не других родственных РНК, находящихся в клетках. В результате в настоящее время считается признанным, что принцип действия рибозимов подтвержден на клеточном уровне.

Для доказательства эффективности действия рибозимов *in vivo* с помощью этих агентов была предпринята попытка индуцировать диабет у мышей путем воздействия на мРНК глюкокиназы β -клеток поджелудочной железы с помощью специфически экспрессирующегося в них трансгена рибозима. Оказалось, что у экспериментальных мышей активность глюкокиназы снижена в три раза именно в β -клетках, что приводило к нарушению секреции ими инсулина и ответа на глюкозу. При этом рибозим не оказывал влияния на функционирование клеток печени. Тем не менее уровень глюкозы в крови таких трансгенных мышей оставался прежним, что позволило сделать вывод о важности дефицита по глюкокиназе печени, а не только поджелудочной железы для развития симптомов заболевания. Эти опыты продемонстрировали возможность использования рибозимов для исследования физиологии у интактных животных.

Эффектные результаты с рибозимами удалось получить и на дрозофиле. Были созданы трансгенные яйца мух, содержащие ген рибозима, действующего на мРНК гена *fushi tarazu* (*ftz*), находящегося под контролем температурно-чувствительного промотора. Этот гомеозисный ген уже упоминался в связи с обсуждением регулирующих его экспрессию факторов транскрипции в первой части книги. Индуцируя рибозим в определенные фазы эмбрионального развития дрозофилы, установили, что вначале нарушение функционирования гена *ftz* приводит к нарушению сегментации кутикулы у личинок. Если ген

инактивировали во время нейрогенеза, то нарушалось развитие центральной нервной системы на фоне нормального развития кутикулы. Эти опыты показали, что индукция рибозима у дрозофилы фенотипически проявляется точно так же, как мутации, инактивирующие ген *ftz*, и не сопровождается другими видимыми нарушениями в организме.

Введение в мужские пронуклеусы оплодотворенных ооцитов мышей плазмиды, экспрессирующей рибозим, направленный против мРНК β_2 -макроглобулина, сопровождалось его последующим синтезом в легких, почках и селезенке взрослых животных. При этом в легких происходило снижение уровня мРНК β_2 -макроглобулина более чем на 90%. Не столь значительное уменьшение содержания мРНК наблюдали в почках и селезенке.

Первые опыты с рибозимами, включенными в состав антисмысловых РНК, показали их высокую эффективность и специфичность *in vitro* и меньшую эффективность *in vivo*. Это может быть связано с тем, что РНК в эукариотических клетках, как правило, находится в составе больших рибонуклеопротеидных комплексов и может быть труднодоступной для антисмысловых РНК, в связи с чем потребуется разработка дополнительных условий для повышения эффективности рибозимов, включенных в состав антисмысловых РНК.

9.2.3. Рибозимы как лекарственные средства

На основании результатов рассмотренных опытов, а также других накопленных знаний о рибозимах складывалось мнение о принципиальной возможности использования рибозимов для регуляции активности конкретных генов человека. Область применения рибозимов как лекарственных средств может быть весьма широкой. В качестве возможных сфер их использования в клиниках в настоящее время рассматривают лечение вирусных инфекций, онкологических заболеваний, для которых известны продукты онкогенов, а также патологических состояний организма, вызываемых сверхэкспрессией конкретных генов.

Лечение вирусных заболеваний. Большие усилия предпринимаются для использования рибозимов в лечении СПИДа, вызываемого вирусами иммунодефицита человека (ВИЧ). Уже в 1990 г. Н. Сарвер и соавторы

продемонстрировали возможность расщепления геномной ВИЧ-РНК с помощью рибозима *in vitro* в бесклеточной системе, что сопровождалось снижением уровня р24-антигена на 98%. Длительные эксперименты на культурах клеток показали, что клетки, экспрессирующие такие рибозимы, не изменяли своего обычного фенотипа на протяжении, по крайней мере, 9 месяцев. Одновременно с этим было показано, что клетки линии МТ₄, стабильно трансфицированные вектором, экспрессирующим рибозим под контролем промотора гена тимидинкиназы вируса простого герпеса, были частично устойчивы к ВИЧ-инфекции. Полностью устойчивыми к заражению ВИЧ оказались клетки, содержащие рибозим под контролем вирусного регуляторного элемента *tar*, в которых вирус не обнаруживали даже через 22 дня после заражения. В этих клетках высокий уровень индукции синтеза рибозима происходит как раз во время ВИЧ-инфекции. Та же самая стратегия была применена и к рибозимам, образующим структуру типа шпильки. Экспрессия гена рибозима под контролем промотора гена β -актина подавляла ВИЧ-инфекцию в культивируемых клетках. При этом продукты расщепления ВИЧ-РНК в таких клетках специфически и быстро деградировали.

Еще одним подходом к преодолению ВИЧ-инфекции является отщепление 5'-концевой кэп-структуры ВИЧ-РНК с помощью рибозимов. Известно, что у большинства изолятов ВИЧ 5'-концевая лидерная последовательность их геномной РНК высококонсервативна. Из всех известных изолятов только штамм MN содержит замену одного нуклеотида в лидерной области РНК, расщепляемой рибозимом. В то же время исследованные штаммы ВИЧ различаются по другим участкам 5'-концевой лидерной последовательности. Эндонуклеазное удаление в составе олигонуклеотида 5'-концевой кэп-группы РНК ВИЧ подавляет трансляцию РНК и делает ее более доступной мишенью для внутриклеточных нуклеаз. Все имеющиеся экспериментальные данные позволяют сегодня сделать вывод о перспективности использования 5'-концевой лидерной последовательности ВИЧ-РНК в качестве мишени для рибозимов.

Основной сложностью, которую приходится преодолевать при конструировании рибозимов, расщепляющих геномную ВИЧ-РНК, является высокая мутабельность этих РНК. Из-за мутационных замен отдельных нуклеотидов в сайтах, по которым происходит расщепление РНК рибозимами,

последние утрачивают свою активность по отношению к этим РНК и соответственно противовирусную активность. Новейшая стратегия в решении данной проблемы заключается в создании нескольких рибозимов, действие которых направлено против разных консервативных участков генома ВИЧ, причем отдельные рибозимы должны входить в одну молекулу рекомбинантной РНК ("метод дробовика"). Помимо последовательностей антивирусных рибозимов такие рекомбинантные РНК содержат в себе дополнительные *cis*-действующие рибозимы, которые осуществляют аутокаталитический процессинг предшественников антивирусных рибозимов с освобождением зрелых рибозимов, действующих на разные участки вирусного генома. Этот метод не позволяет вирусу ВИЧ избежать инактивирующего действия рибозимов путем мутационных преобразований своего генома.

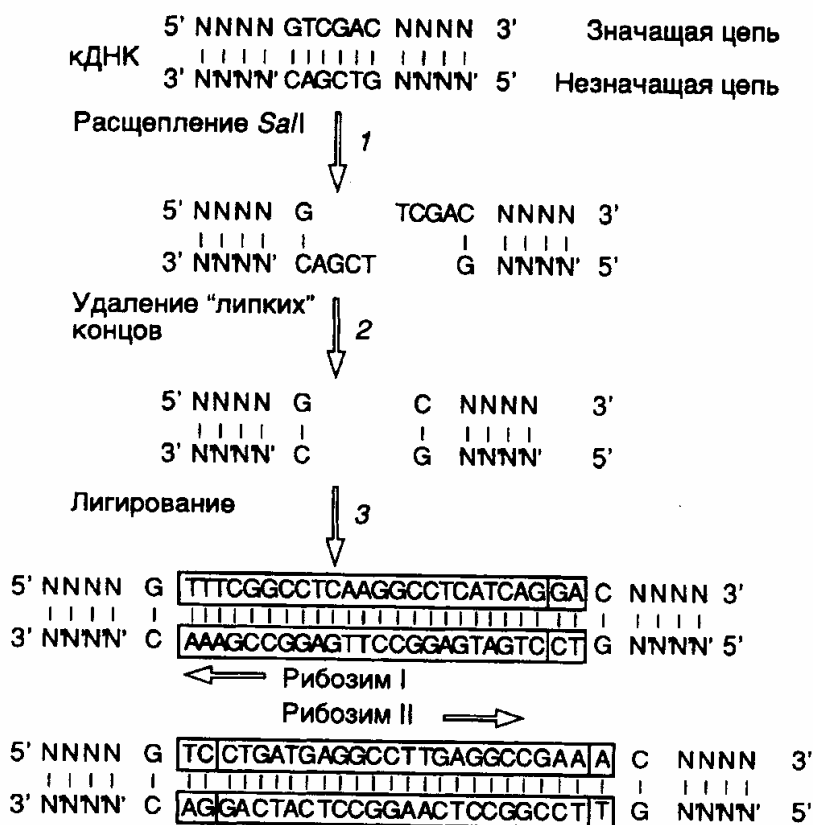


Рис. II.26. Использование сайтов рестрикции для конструирования рибозима, интегрированного в антисмысловую РНК

Недавно была получена кассета из девяти идентичных генов рибозимов, расположенных друг за другом в одном экспрессирующем векторе. Рибозимы такой "полицистронной" РНК, направленные против разных частей ВИЧ-генома, сохраняли внутри клетки свою специфичность, но были более эффективными в пересчете на одну копию РНК полирибозима и не утрачивали своей активности даже в частично деградировавших молекулах РНК.

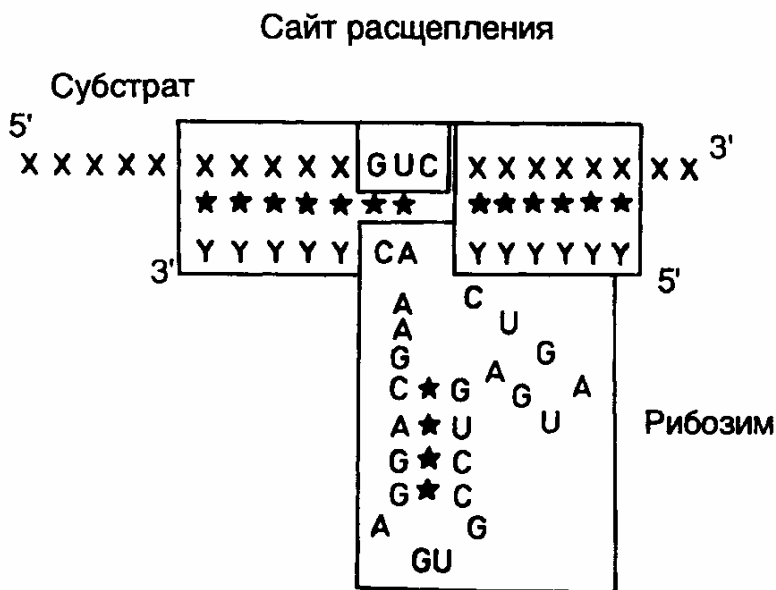


Рис. II.27. Взаимодействие рибозима с РНК-субстратом

В заключение этого раздела рассмотрим пример использования рибозимов против вирусов растений в сельском хозяйстве. Как упоминалось выше, рибозимы, образующие структуру типа "головки молотка", распознают в сайте расщепления РНК-мишени последовательности из трех нуклеотидов, рядом с которой происходит гидролиз цепи РНК. Подобными тринуклеотидами могут быть последовательности AUC, CUA, CUC, GUA, GUC, GUU и UUC. В предложенном подходе РНК-связывающий домен рибозима конструируют таким образом, что распознаваемый тринуклеотид входит в состав последовательности нуклеотидов, которая в двухцепочечной ДНК образует сайт рестрикции. Подобный подход дает возможность встраивать последовательность каталитического домена рибозима по этому сайту рестрикции в кДНК, полученную путем обратной транскрипции РНК-мишени.

Следовательно, каталитический домен рибозима оказывается фланкированным последовательностями антисмысловой РНК, комплементарной РНК-мишени, которые автоматически выполняют роль РНК-связывающего домена. При таком подходе отпадает необходимость в синтезе большого числа олигонуклеотидов, и в любом необходимом случае рибозим собирается из готовых блоков. В качестве конкретного примера рассмотрим создание рекомбинантного рибозима, специфически расщепляющего РНК, содержащую последовательность сайта рестрикции *SaII* (рис. II.26). После расщепления кДНК рестриктазой *SaII* (этап 1) образовавшиеся "липкие" концы гидролизуют нуклеазой S1 (этап 2) и по образовавшимся "тупым" концам встраивают каталитический домен рибозима, заключенный в двухцепочечном фрагменте ДНК (этап 3). В результате транскрибирования такой конструкции синтезируется антисмысловая РНК, в состав которой включен каталитический домен рибозима. После гибридизации с РНК-мишенью образуется двухцепочечная молекула РНК, на одной из цепей которой формируется структура рибозима типа "головки молотка" (рис. II.27), расщепляющего РНК-мишень в строго определенном месте. Описанный подход был реализован при конструировании рибозима, направленного против геномной РНК одного из вирусов сливы. При этом были получены последовательности нуклеотидов рибозима, пригодные для использования с другими 25 сайтами рестрикции.

Подавление функционирования онкогенов. Одними из основных различий, обнаруживаемых между нормальными и раковыми клетками, являются генетические различия ряда генов, контролирующих пролиферацию. В геноме опухолевых клеток часто обнаруживают мутации в генах двух типов: онкогенах и генах-супрессорах опухолевого роста, или антионкогенах. Онкогены эволюционно консервативны и вызывают неопластическую трансформацию клеток как при ретровирусной инфекции в природных условиях (поскольку они часто бывают включены в состав генома ретровирусов), так и после введения ДНК онкогенов в культивируемые клетки с помощью трансфекции. Большинство онкогенов вначале было обнаружено именно в составе генома онкогенных вирусов, и они являются мутантными производными протоонкогенов, присутствующих в здоровых клетках многоклеточных организмов и активирующихся во время эмбриогенеза, роста клеток или регенерации тканей. Поскольку активированные онкогены в опухолевых клетках, как правило,

сверхэкспрессируются и кодируемая ими РНК по своей первичной структуре отличается от РНК протоонкогенов, РНК онкогенов являются хорошей потенциальной мишенью для рибозимов.

Онкоген *H-ras*. Мутация в кодоне 12 гена *H-ras*, приводящая к замене GGC→GUC, создает консенсусный сайт, по которому НН-рибозим может расщеплять мутантную мРНК. In vitro было продемонстрировано пятикратное различие в эффективности действия рибозима на мутантную *H-ras*-РНК и соответствующую РНК дикого типа. Получены *H-ras*-зависимые линии клеток, стабильно трансформированные экспрессирующим вектором, который направлял синтез НН-рибозима под контролем промотора β-актинового гена. Для таких клеток характерна пониженная скорость пролиферации, сопряженная с уменьшением внутриклеточных уровней *H-ras*-РНК и белка p21, кодируемого этим геном. Далее рибозим экспрессировали в клетках линии EJ карциномы мочевого пузыря человека. Введение исходных клеток мышам сопровождалось их быстрой гибелью на фоне развития высокоинвазивных опухолей. В отличие от этого клоны EJ-клеток, экспрессирующих рибозим, в организме мышей обладали резко сниженным опухолевым фенотипом. Образующиеся опухоли были малоинвазивны, и наблюдалось приблизительно двукратное повышение уровня выживаемости мышей с трансплантатами. Гистологические исследования подтверждали слабую способность опухолей к метастазированию. Рибозимы в опухолях обнаруживались методом ПЦР на протяжении 86–90 дней. Таким образом, эти и ряд других экспериментов показали способность рибозимов нормализовать экспрессию активированного *H-ras* как in vitro, так и in vivo.

Онкоген *c-fos*. Как уже упоминалось в первой части книги, белковый продукт гена *c-fos* участвует в передаче сигнала эукариотическими клетками, вовлечен в синтез ДНК и может придавать клеткам устойчивость к противоопухолевым препаратам. Последние два свойства этого белка находят подтверждение, в частности в том факте, что при проведении лечения часто используемым противоопухолевым лекарством цисплатином (цис-диаминодихлорплатина) происходит индукция гена *c-fos* вслед за генами dTMP-синтазы и ДНК-полимеразы. Рибозим, разрушающий *c-fos*-мРНК, снижал конечный уровень экспрессии гена *c-fos*, приводил к повышению чувствительности опухолевых клеток к химиотерапевтическим агентам

(включая *цис*-платину) и значительно подавлял экспрессию генов dTMP-синтазы, ДНК-полимеразы β и гена металлотioneина человека hMTII-A.

Онкогены *bcr/abl*. Аберрантная филадельфийская хромосома образуется в результате транслокации (9;22)(q34;q11) в стволовых клетках костного мозга, что сопровождается слиянием генов *bcr* и *abl* с образованием химерного онкогена *bcr/abl* и развитием хронических миелоидных лейкозов (ХМЛ). Транскрипт химерного гена кодирует белок p210^{*bcr/abl*}, который обладает повышенной активностью тирозинкиназы. Такие РНК и белок обнаруживаются почти у всех больных с синдромом ХМЛ, а также у 50% пациентов с острым лимфобластическим лейкозом, у которых имеется филадельфийская хромосома. С помощью рибозимов, специфически расщепляющих последовательность химерной мРНК в месте стыковки последовательностей двух генов, удалось подавить экспрессию химерного гена *bcr/abl* в культивируемых клетках. Экспрессия рибозимов в культивируемых клетках вызывала снижение уровня *bcr/abl*-мРНК, полностью блокировала образование химерного белка p210^{*bcr/abl*} и ингибировала рост клеток на 84%. Эти результаты были значительно лучше эффектов, вызываемых антисмысловыми олигонуклеотидами.

В работах других исследователей был получен рибозим, расщепляющий вышеупомянутую химерную РНК по кодону GUU, расположенному по соседству с местом стыковки последовательностей двух генов. Этот рибозим разрушал соответствующую РНК *in vitro* и *in vivo*, а также подавлял тирозинкиназную активность белка p210.

9.2.4. Репарация мутантных РНК с помощью рибозимов, осуществляющих *транс-сплайсинг*

Просто устроенные НН- и НР-рибозимы принято рассматривать как рибозимы первого поколения. Во всех примерах, рассмотренных выше, в прикладных целях использовалась их способность осуществлять специфическое эндонуклеазное расщепление молекул РНК. Недавно (Дж. Т. Джонс и др., 1996 г.) удалось приспособить рибозимы для восстановления генетической информации на уровне молекул РНК, структура которых нарушена мутациями. Рибозим интрона группы I *Tetrahymena* после

незначительных генно-инженерных модификаций оказался способным выполнять такие репарирующие функции.

Как уже обсуждалось в первой части книги (см. раздел 2.2.3), аутосплайсинг транскриптов, содержащих интроны группы I, начинается с расщепления фосфодиэфирной связи между 5'-концевым экзоном и прилегающим к нему интроном, сопровождаемого полным удалением интрона и лигированием фланкирующих его экзонов, приводящим к восстановлению непрерывной структуры РНК. Идея, лежащая в основе использования рибозимов для репарации РНК, проста. Достаточно лишь с помощью генно-инженерных методов к 3'-концу рибозима присоединить вместо 3'-концевого экзона любую другую последовательность, чтобы она соединялась с последовательностью 5'-концевого экзона в строго определенной точке с образованием гибридной молекулы РНК. Поскольку сайт взаимодействия рибозима с РНК во время *транс*-сплайсинга и положение точки разрыва РНК целиком определяются внутренней направляющей последовательностью рибозима (internal guide sequence – IGS), присоединение новой 3'-концевой последовательности можно осуществлять практически в любом месте РНК-мишени (рис. II.28). Таким образом, для репарации мутационного повреждения в экзоне 2 на уровне РНК IGS делают комплементарной участку РНК-мишени, локализованному выше точки повреждения (в рассматриваемом примере – А) на стыке экзонов 1 и 2. Экзон 2 дикого типа соединяют с рибозимом, который после взаимодействия с РНК-мишенью вырезает из нее мутантную последовательность экзона 2 и заменяет на эквивалентную последовательность дикого типа. В результате комплекс рибозима и репарированной РНК распадается и обновленная РНК вовлекается в трансляцию.

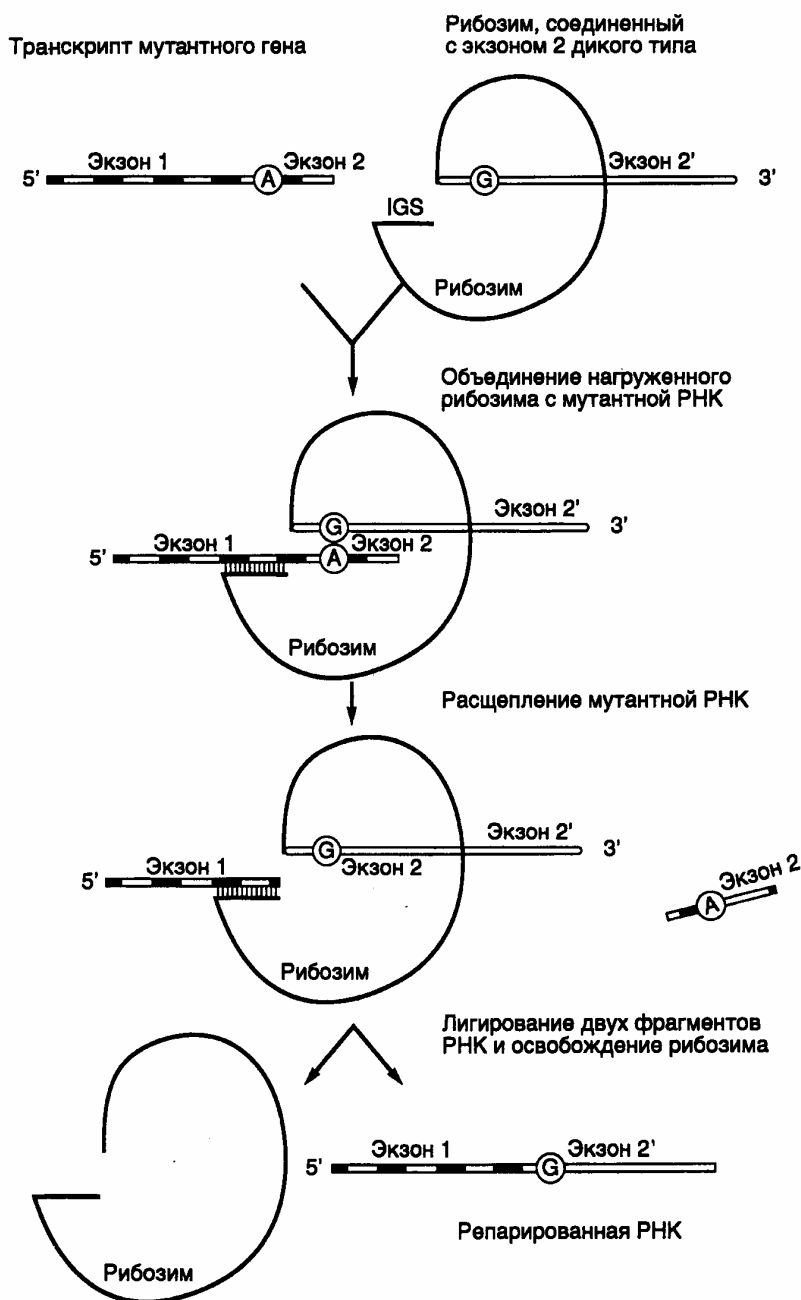


Рис. II.28. Схема репарации мутантной мРНК рибозимом, осуществляющим *транс*-сплайсинг

Молекула рибозима, соединенная 3'-концом с последовательностью экзона 2 дикого типа, взаимодействует с мутантной РНК выше мутантного нуклеотида и замещает мутантную часть репарируемой РНК в реакции аутосплайсинга

Такой подход был успешно использован для репарации мРНК гена *lacZ* как в клетках *E. coli*, так и в культивируемых клетках животных. Точность *транс*-сплайсинга, направляемого рибозимом, оказалась абсолютной – в образующихся восстановленных молекулах РНК не наблюдали сдвига рамки считывания. Однако в этих опытах на первый план вышла другая проблема. Длина IGS-последовательности, определяющей специфичность взаимодействия рибозима с РНК, составляет всего шесть нуклеотидов. Такая последовательность из шести нуклеотидов в молекулах РНК встречается достаточно часто (каждые 4096 оснований). Следствием этого было объединение репарирующего фрагмента РНК не только с РНК-мишенью, но и другими РНК, не имеющими отношения к первой. Выход из сложившейся ситуации ищут в увеличении специфичности действия рибозима путем удлинения IGS. Однако с увеличением длины области комплементарного взаимодействия между рибозимом и РНК будут возрастать прочность этого взаимодействия и, как следствие, затруднение последующего распада каталитического комплекса, что должно сопровождаться ингибированием реакции.

Применение рибозимов в реакциях *транс*-сплайсинга для корректировки экспрессии мутантных генов на посттранскрипционном уровне рассматривается как весьма перспективное направление в генотерапии. И хотя более логичной все-таки кажется необходимость репарации мутаций на уровне самих генов, такой подход, возможно, получит право на жизнь и в клиниках. По сути дела он является прототипом реакции редактирования РНК (см. раздел 2.2.2), который эффективно используется клетками при экспрессии ряда генов в природных условиях.

Благодаря генной инженерии введение в молекулярную генетику антисмысловых РНК и рибозимов привело к появлению высокоспецифичного инструмента для изучения функциональной роли и регуляции экспрессии индивидуальных генов *in vivo*. Использование этих методов представляется особенно перспективным в исследованиях по биологии и генетике развития, так как именно здесь инактивация отдельных генов соматических клеток представляется особенно сложной задачей. Плодотворным подходом является объединение последовательностей антисмысловых РНК с рибозимами, поскольку в результате получают каталитически действующие,

высокоспецифические супрессоры генов, а также конструкции, корректирующие экспрессию генов на посттранскрипционном уровне. Полагают, что дальнейшее развитие техники адресной доставки генов антисмысловых РНК, направленных против патогенных вирусов, в клетки зародышевой линии животных и растений может привести к созданию сортов растений и пород животных, устойчивых к действию патогенных вирусов, и дать большой экономический эффект в сельском хозяйстве. Однако хорошо известно, что на любое лекарство у возбудителя заболевания всегда находится адекватный защитный ответ. Поэтому до тех пор, пока генотерапия не будет опираться на самонастраивающиеся генетические системы, что имеет место, например в случае иммунной системы, победа, в конечном счете, будет оставаться за возбудителем. Нельзя не учитывать и возможное использование антисмысловых РНК в будущем для генотерапии как инструмента воздействия на онкогены и другие генные локусы, повышенный уровень экспрессии которых вредно сказывается на жизнедеятельности соматических клеток многоклеточных организмов. Первые результаты, полученные при разработке некоторых из этих направлений молекулярной генетики, кратко изложены в главе 11.

9.2.5. Дезоксирибозимы

В отличие от РНК, выполняющих в клетке разнообразные функции, благодаря возможностям формирования у этих макромолекул сложных пространственных структур, для внутриклеточных ДНК пока известна единственная функция – пассивное хранение генетической информации. Отчетливо различающиеся биохимические функции биополимеров этих двух видов объясняют исключительно особенностями их химического состава. Отсутствие у ДНК в 2'-положении рибозы гидроксильной группы делает ее более устойчивой к расщеплению *in vivo*, а следовательно, более приемлемым, чем РНК, носителем генетической информации. Наличие же у РНК 2'-ОН-группы, напротив, является предпосылкой ее структурной и функциональной сложности.

Несмотря на прогноз о большой каталитической инертности ДНК, в течение последних пяти лет ферментативная активность была обнаружена и у ее молекул, правда пока только в искусственных условиях. Молекулы ДНК,

обладающие ферментативной активностью, получили название *дезоксирибозимов*. К настоящему времени описана способность дезоксирибозимов расщеплять, лигировать и фосфорилировать молекулы ДНК, метилировать порфирины или осуществлять гидролиз РНК. Получены дезоксирибозимы с активностями РНКазы, которые используют в качестве кофакторов аминокислоты, ионы металлов, а также не требуют кофакторов вообще. При этом Mg^{2+} -зависимые дезоксирибозимы расщепляют РНК со скоростью, сопоставимой с таковой природных рибозимов ($k_{кат} = 10 \text{ мин}^{-1}$).

Поиск молекул ДНК, обладающих определенными ферментативными активностями, как правило, осуществляется с применением скрининга пула молекул со случайной первичной структурой. Результативность таких исследований во многом зависит от эффективности системы скрининга. Так, при отборе дезоксирибозимов с полинуклеотидкиназной активностью И. Ли и Р.Р. Брикер (1999 г.) использовали пул из 10^{15} различных молекул синтетических одноцепочечных ДНК длиной в 100 нуклеотидов со случайной 70-нуклеотидной последовательностью в центральной части. Из пула отбирали молекулы, способные фосфорилировать свой 5'-конец в присутствии различных рибо- или дезоксирибонуклеозидтрифосфатов. Наличие 5'-концевого фосфата обнаруживали по приобретению олигонуклеотидами способности участвовать в реакции лигирования с акцепторным олигонуклеотидом на одноцепочечной матрице, стыкующей их концы, поскольку только фосфорилированные молекулы ДНК могли лигироваться с акцепторной ДНК. Молекулы ДНК, длина которых увеличивалась в результате лигирования, очищали электрофорезом, амплифицировали и использовали в новых раундах скрининга. В итоге из смеси молекул ДНК были выделены олигонуклеотиды, фосфорилирующие свой 5'-концевой нуклеотид с использованием γ -фосфатной группы рибо- и дезоксирибонуклеотидов. Некоторые олигонуклеотиды оказались высокоспецифичными в отношении доноров фосфатных групп и работали только в присутствии конкретных нуклеотидов, тогда как другие были неразборчивы в отношении индивидуальных донорных нуклеотидов. Более того, структурные исследования показали, что полученные молекулы ДНК обладают достаточной сложностью, чтобы делать выбор между различными молекулами субстратов, кофакторов и эффекторов.

На основании такого рода данных были сделаны выводы, что молекулы ДНК, как и РНК, способны формировать различные третичные структуры, необходимые для осуществления химического катализа. В будущем они могут быть использованы для конструирования различных дезоксирибозимов. При этом не исключается, что дезоксирибозимы могут быть обнаружены и *in vivo*.

9.3. Аптамеры

Аптамерами называют небольшие молекулы нуклеиновых кислот, которые могут выполнять функции высокоспецифичных рецепторов низкомолекулярных органических соединений. Олигонуклеотидные аптамеры с требуемыми свойствами выделяют из библиотек случайных последовательностей методами селекции *in vitro*, используя их способность специфически взаимодействовать с соответствующими иммобилизованными лигандами. Такие соединения находят применение в фундаментальных исследованиях молекулярных механизмов взаимодействия нуклеиновых кислот с лигандами, и их начинают использовать для обнаружения повреждений в ДНК, для воздействия на экспрессию генов, а также в качестве модулей при конструировании рибозимов, обладающих аллостерическими свойствами. В настоящее время получены аптамеры как рибо-, так и дезоксирибонуклеотидной природы, которые образуют специфические комплексы с разнообразными химическими соединениями, включая большинство аминокислот, нуклеотиды и их производные, антибиотики, биологически активные пептиды, органические красители, дофамин, теofilлин, и многие другие органические соединения, имеющие биологическое значение. Приведем несколько примеров, иллюстрирующих последние достижения в этой области исследований.

Аптамеры, взаимодействующие с красителями. Недавно была реализована идея использования аптамеров для введения флуоресцентной метки в молекулы РНК, синтезируемые *in vivo*, для наблюдения за экспрессией соответствующих генов. Из пула случайных последовательностей были выделены ~60-звенные олигорибонуклеотиды, специфически взаимодействующие с сульфородамино В. После введения в аптамеры мутаций и повторной селекции удалось получить олигонуклеотиды, взаимодействующие с флуоресцеином. В иммобилизованном состоянии эти

аптамеры специфически связывали флуоресцентные красители и не давали перекрестных реакций.

В другой серии исследований были идентифицированы аптамеры, взаимодействующие с красителем Hoechst H33258, которые функционировали как *in vitro*, так и *in vivo*. Последовательности аптамеров в форме кДНК встраивали в виде tandemных повторов в нетранслируемую область гена β -галактозидазы, находящегося в составе экспрессирующего вектора. Вектор вводили с помощью трансфекции в клетки линии СНО, которые далее выращивали в присутствии красителя. При определении активности β -галактозидазы через 24 ч после трансфекции установили, что ее синтез не происходил если Hoechst H33258 присутствовал в питательной среде в концентрации 5–10 мМ.

Аптамеры, взаимодействующие с небольшими молекулами биологического происхождения. В качестве первого этапа серии исследований по созданию аптамеров, которые бы специфически распознавали олигосахариды поверхности клеток, недавно (1998 г.) были получены олигодезоксирибонуклеотиды, специфичные в отношении дисахарида целлобиозы. Для скрининга аптамеров использовали простую и эффективную методику. Библиотеку олигонуклеотидов инкубировали с целлюлозой, цепи молекул которой построены из целлобиозы, и несвязавшиеся олигонуклеотиды отмывали буферным раствором, а оставшиеся в комплексе — элюировали раствором целлобиозы. Оказалось, что полученные аптамеры специфически распознают целлобиозу как в растворе, так и в составе целлюлозы ($K_d = 10^{-0,1}$ мкМ) и не взаимодействуют с родственными дисахаридами: мальтозой, лактозой и гентиобиозой. Такие результаты создают предпосылки для получения полного набора олигонуклеотидных рецепторов, который можно было бы использовать в технологии микроматриц ДНК для идентификации антигенов поверхности клеток, что весьма актуально, в частности для осуществления пересадки органов и тканей.

Другим впечатляющим результатом этого направления исследований было недавнее получение аптамеров, которые обладали в 1000 раз большим сродством к *N*-7-метилгуанозину (основному компоненту кэп-групп мРНК эукариот), чем к его неметилированному аналогу. Такие аптамеры специфически ингибировали трансляцию мРНК, содержащих кэп-группы, в

бесклеточных системах, полученных из клеток HeLa или дрожжей. Полагают, что такие аптамеры найдут применение в исследовании кэп-зависимых процессов метаболизма эукариотических мРНК, например сплайсинга пре-мРНК или их транспорта из ядра в цитоплазму.

С помощью аптамеров, специфичных в отношении 7,8-дигидро-8-гидрокси-2'-дезоксигуанозина, который возникает при повреждении ДНК радикалами кислорода, удавалось обнаруживать в ДНК одну такую молекулу среди 10^4 обычных остатков дезоксигуанозина. Аналогичные аптамеры могут быть потенциально использованы для мониторинга мутагенных последствий окислительного стресса в эукариотических клетках.

Перспективы использования аптамеров. Новейшие результаты, полученные с использованием аптамеров, указывают на возможность их применения в ближайшем будущем в биотехнологии для регулируемой экспрессии генов, а также для диагностики и терапии. Одним из наиболее интересных достижений в этой области исследований является создание рибозимов с аллостерическими свойствами. В таких конструкциях каталитический домен рибозима соединяют с последовательностями аптамеров-рецепторов, взаимодействующих с регуляторными молекулами. В присутствии низкомолекулярных регуляторов может происходить специфическая активация или ингибирование ферментативной активности рибозима. Такие гибридные конструкции получили название *аптазимов*. Недавно полученный аптазим, обладающий РНК-лигазной активностью, активировался в 10^5 раз в присутствии специфических низкомолекулярных эффекторов. При аналогичном подходе были сконструированы АТФ-зависимые рибозимы.

Исключительно важные результаты могут быть получены при скрининге *in vitro* аптамеров среди пула фрагментов природных РНК. В этом случае библиотеки природных последовательностей могут быть использованы для поиска доменов нуклеиновых кислот, специфически взаимодействующих с известными молекулами биогенного происхождения, например белками ретровирусов, для которых неизвестны РНК-мишени клетки-хозяина. Не исключено, что в ходе такого рода исследований у РНК или их фрагментов будут обнаружены новые внутриклеточные функции, связанные с

метаболизмом важных биологических кофакторов, включая ATP, GTP, FAD и NAD⁺.

9.4. Молекулы РНК у истоков жизни

Большинство современных теорий происхождения жизни рассматривает молекулы РНК, обладающие активностями рибозимов, в качестве первичных самореплицирующихся молекул, давших начало развитию жизни на Земле. Результаты исследования двух гипотетических групп рибозимов лежат в основе современных представлений о происхождении жизни. Прежде всего, это РНК-репликазы, построенные исключительно из рибонуклеотидов, т.е. молекулы РНК, способные самореплицироваться. Такое предположение кажется особенно привлекательным, поскольку древние молекулы РНК, с одной стороны, могли бы хранить генетическую информацию, а с другой – обладать ферментативной активностью, необходимой для воспроизведения этой информации. Подобное сочетание свойств в молекуле РНК позволяет объяснить, каким образом каталитические функции, выполняемые современными белковыми молекулами, а также функции хранения и передачи генетической информации возникли одновременно и могли присутствовать в одной молекуле-предшественнике прародителей современных клеток. При переходе от первоначального мира самореплицирующихся молекул РНК к современному живому миру, повсеместно использующему белковый катализ, предполагают возникновение молекул рибозимов, способных осуществлять синтез белка из абиотически возникших его предшественников – аминокислот.

Теории происхождения жизни, основанные на первоначальном участии молекул РНК в создании самореплицирующихся и самоусложняющихся систем, нашли прочную поддержку в связи с открытием аутосплайсинга и рибозимов. Как будет видно из дальнейшего изложения, механизм аутосплайсинга не сильно отличается от комплекса химических реакций, необходимых для репликации нуклеиновых кислот. Кроме того, в настоящее время все больше утверждается точка зрения, в соответствии с которой именно современные рибосомные РНК составляют каталитическую сердцевину рибосомы и могут быть потомками рибозимов, самостоятельно осуществлявших синтез полипептидов. В этой связи представляют интерес некоторые гипотетические схемы молекулярных механизмов экспрессии предшественников современных

генов.

9.4.1. Молекулы РНК в качестве РНК-репликазы

Для первоначального появления рибозимов необходим абиотический синтез олигорибонуклеотидов длиной 30–70 оснований. Долгое время это требование было камнем преткновения в разработке теории происхождения жизни, основанной на РНК. Недавно было показано, что после спонтанной иммобилизации на твердой поверхности коротких олигонуклеотидов в присутствии свежих порций активированных мономеров их длина может самопроизвольно удлиниться до 55 остатков. Значительным осложнением для обоснования теории было также то, что мономеры в спонтанно образующихся олигонуклеотидах соединены друг с другом смешанными 2'–5'- и 3'–5'-связями. Такая гетерогенность олигонуклеотидов должна затруднять их последующую матричную активность. Оказалось, однако, что олигоцитидилаты, содержащие смешанные связи, могут служить матрицами для спонтанной полимеризации мономеров, активированных имидазолом, с образованием олигогуанилатов. Таким образом, вышеупомянутые возражения в настоящее время едва ли существенны для обсуждаемой теории.

Для того чтобы репликация РНК самими молекулами РНК имела место, необходимо одновременное присутствие в одном микрообъеме, по крайней мере, двух молекул рибозима-репликазы или же репликазы и комплементарной ей цепи РНК, которая служила бы матрицей для репликазы. В таком случае цикл воспроизводства молекулы репликазы можно представить следующим образом. После разделения цепей синтезированной репликазы и матрицы с последней соединяется олигонуклеотидный праймер, который последовательно удлиняется рибозимом-репликазой, образовавшейся в результате фолдинга второй цепи. Предполагается, что неорганические поверхности, обладающие большим сродством к одноцепочечным РНК, способствуют разделению цепей после завершения акта репликации. При этом должна существовать определенная компартментализация таких молекул (например обеспечиваемая силами адсорбции), препятствующая слишком далекому расхождению разделившихся цепей РНК. И, наконец, синтез РНК, осуществляемый древней репликазой, должен характеризоваться определенным уровнем точности.

Выше уже были рассмотрены механизмы реакций аутосплайсинга и *транс*-сплайсинга, осуществляемые рибозимами интронов группы I Tetrahymena. В основе этих процессов лежит основанная на переэтерификации аутокаталитическая реакция объединения (лигирования) двух полирибонуклеотидных последовательностей (экзонов), сопровождающаяся удлинением акцепторной последовательности, которое можно рассматривать как прототип реакции полимеризации. В настоящее время экспериментально подтверждена возможность участия в таких реакциях более коротких олигонуклеотидов. В частности, благодаря этим реакциям пентамеры цитидиловой кислоты (C_5) превращались в гетерогенную популяцию олиго(C), содержащую даже 30-членные олигонуклеотиды. Вскоре было показано, что и динуклеотиды GpC могут выступать в качестве доноров мононуклеотидов при элонгации C_5 с образованием C_{10} в реакции удлинения праймера ($C_5 + 5GpC \rightarrow C_{10} + 5G$). Частота ошибок в такой реакции в условиях насыщения GpC-субстратами составляла ~35% и была слишком большой для того чтобы обеспечивать саморепликацию молекул РНК.

Альтернативой реакции удлинения праймера на один нуклеотид за цикл может быть зависящая от матрицы реакция элонгации праймера олигонуклеотидами, которая также была продемонстрирована экспериментально. Здесь донорами выступали триплеты олигорибонуклеотидов и, как уже упоминалось выше, крайним известным случаем такой реакции является объединение экзонов. При этом рибозимы Tetrahymena в качестве матриц для сборки олигонуклеотидов могут использовать различные внутренние последовательности нуклеотидов. Так, олигонуклеотиды длиной 8–12 остатков недавно были применены для сборки полной цепи, комплементарной последовательности самого рибозима.

Низкий выход продуктов полной длины в экспериментах, упомянутых выше, позволял предполагать, что с этой задачей успешнее могли бы справиться более короткие и активные рибозимы, которые пока еще в природе неизвестны. В связи с этим была разработана система отбора *in vitro*, позволившая повысить эффективность реакции лигирования путем получения более коротких производных известного рибозима *sunY*. Приблизительно $2 \cdot 10^{13}$ вариантов нуклеотидных последовательностей рибозима в виде единого пула

были подвергнуты нескольким циклам все более строгого отбора вариантов рибозима на повышенную способность к лигированию с последующей амплификацией последовательностей в конце каждого цикла отбора. Такие варианты рибозимов с более короткой полинуклеотидной цепью вскоре были получены. Они обладали способностью эффективно осуществлять из 18 олигонуклеотидов полную сборку цепи РНК, комплементарной своей собственной полинуклеотидной цепи. Наиболее интересной находкой в ходе такого рода экспериментов по отбору из пула случайных последовательностей оказался рибозим длиной в 100 нуклеотидов, получивший название *лигазы класса I*. Этот рибозим осуществлял образование 3'–5'-фосфодиэфирных связей, и его эффективность была сопоставима с таковой фермента – природной ДНК-лигазы ($k_{\text{кат}} > 1 \text{ с}^{-1}$). Недавно было продемонстрировано, что лигаза класса I может выполнять функции РНК-полимеразы, осуществляя элонгацию праймера путем использования рибонуклеозидтрифосфатов в качестве субстратов. Основным ограничением этой реакции является то, что поскольку комплекс рибозим–матрица удерживается комплементарными взаимодействиями, элонгация праймера прекращается, как только 3'-конец образующегося продукта достигает конца матрицы. В реакции не только используются все четыре рибонуклеозидтрифосфата, но и достигается большая точность синтеза РНК (92%). Дальнейшее ее повышение хотя бы в 10 раз позволило бы этой искусственной РНК-полимеразе воспроизводить свою собственную последовательность нуклеотидов. Полагают, что полное превращение рибозима в РНК-полимеразу произойдет после того, как матрица будет связываться с ним некомплементарными взаимодействиями.

Условия отбора, использованные для получения эволюционирующих *in vitro* молекул лигазы класса I, были отработаны на современных субстратах (рибонуклеозидтрифосфатах). Не исключено, что изменение этих условий позволит получить рибозимы с новыми активностями. Недавно был описан рибозим, осуществляющий реакцию, аналогичную кэпированию акцепторного олигонуклеотида, которая завершалась присоединением к его 5'-концу молекулы аденозинмонофосфата посредством 5'–5'-фосфодиэфирной связи. Полагают, что, используя в качестве субстратов высокоактивированные рибонуклеозидмонофосфаты, можно будет получить рибозимы, непосредственно полимеризующие мононуклеотиды.

9.4.2. Возможность синтеза полипептидных цепей молекулами РНК

Вскоре после открытия рибозимов в литературе стала активно обсуждаться гипотеза о каталитической (а не только структурной) функции рРНК в рибосомах. Первые экспериментальные данные в пользу возможной каталитической функции рРНК были получены с ингибиторами биосинтеза белка. Оказалось, что устойчивость к некоторым из них вызывается мутациями в генах рРНК, но не рибосомных белков. Х.Ф. Ноллером и соавторами (1992 г.) было установлено, что удаление из рибосомы большого числа рибосомных белков не инактивирует ее пептидилтрансферазную функцию. Подобного рода данные указывают на то, что именно 23S рРНК является пептидилтрансферазой, однако окончательный вывод мешают сделать оставшиеся рибосомные белки, присутствие которых необходимо для проявления пептидилтрансферазной активности такой ослабленной в структурном отношении рибосомы.

Функциональная связь между рРНК и интронами группы I недавно открылась с неожиданной стороны, когда установили, что аминогликозидные антибиотики, блокирующие биосинтез белка, ингибируют и аутосплайсинг интронов группы I. Это открытие нашло дальнейшее развитие в связи с обнаружением у интронов группы I слабой аминокислэстеразной активности. Было сделано предположение, что интроны группы I и рибосомная РНК могут обладать общими структурными (и функциональными) доменами. Кроме того, мутантные рибозимы, полученные путем отбора *in vitro* по критерию ускоренного переноса фосфодиэфирных связей, оказались способными, хотя и в слабой степени, расщеплять амидные связи. На основании такого рода данных Ноллер и соавторы сделали предположение о том, что первые примитивные рибосомы образовались из молекул РНК, родственных интронам группы I, которые были в состоянии осуществлять перенос ацильных групп, необходимый для образования пептидных связей. Таким образом, все эти данные в совокупности позволяли предполагать, что функция рРНК в биосинтезе белка является каталитической, а некоторые РНК сегодняшнего дня, имеющие отношение к обсуждаемым процессам – это "молекулярные ископаемые".

При наличии активированных аминокислот синтез пептидов не

представляется трудной задачей. Активированные аминокислоты конденсируются даже в водных растворах с образованием коротких пептидов, а цепи длиной до 50 аминокислот образуются на минеральных поверхностях. Абстрактная схема биосинтеза белка в примитивных системах с участием каталитических РНК представлялась следующим образом. Примитивные РНК, аминоацилирующие сами себя активированными аминокислотами по аутокаталитическому механизму, могут выступать донорами и акцепторами аминокислот в реакциях переноса ацильных групп, катализируемых рибозимами. Для признания РНК в качестве молекул, осуществлявших в примитивных системах синтез белков, необходимо показать возможность выполнения ими следующих функций: узнавание аминокислот, аминоацилирование тРНК, перенос ацильных групп, активация аминокислот и синтез пептидов. Рассмотрим реальность каждого из этих предположений.

Молекулы РНК могут распознавать аминокислоты. По крайней мере, у двух природных РНК (как и у обсуждавшихся выше искусственных аптамеров) обнаружены структурные элементы, способные связывать аминокислоты. Интроны группы I связывают аргинин в области спирали Р7 (см. рис. I.15,а). Вторым примером является трансактивируемый район (*trans*-activating region – TAR) ВИЧ, который специфически взаимодействует с белком-трансактиватором tat во время вирусной инфекции. Этот участок ВИЧ-РНК также обладает способностью взаимодействовать с аргинином. Кроме того, известно, что редактирующая функция аминоацил-тРНК-синтетаз (разрушение связи между ошибочно соединенными аминокислотой и тРНК) включает опосредованное РНК распознавание аминокислот. Предполагается, что в процессе редактирования имеют место непосредственные контакты РНК с аминокислотами. Более того, в системах с искусственным отбором РНК из пула молекул со случайными последовательностями удалось получить четыре различных аргининсвязывающих структурных элемента (аптамера) с разным уровнем специфичности. С помощью того же подхода были выделены акцептирующие цитруллин аптамеры, а также другие молекулы РНК, взаимодействующие с более гидрофобными аминокислотами – валином и триптофаном.

Спонтанное аминоацилирование РНК. Процесс аминоацилирования тРНК аминоацил-тРНК-синтетазами *in vivo* происходит в два этапа: вначале

фермент активирует соответствующую аминокислоту с образованием 5'-аминоациладенилата, который затем атакуется 2'(3')-концом тРНК, что сопровождается возникновением соответствующей ковалентной связи между аминокислотой и тРНК. Нагруженные таким образом молекулы тРНК далее используются рибосомами в качестве субстратов при образовании пептидных связей. Недавно тем же методом селекции из пула случайных РНК, представленного $1,7 \cdot 10^{14}$ молекулами, были получены рибозимы, способные при наличии активированных аминокислот спонтанно аминоацилироваться. Пул РНК инкубировали при 0° с химически синтезированным фенилаланил-5'-аденилатом в присутствии ионов Mg^{2+} и Ca^{2+} . Аминоцилированные РНК обнаруживали по появлению свободной α -аминогруппы фенилаланина. Молекулы РНК, меченные такими гидрофобными группами, очищали с помощью высокоэффективной жидкостной хроматографии (ВЭЖХ) в обращенной фазе. После 11 циклов отбора были получены спонтанно аминоацилирующиеся РНК со скоростью, в 105 раз превышающей скорость реакции в контроле. Распознавание аминокислоты рибозимом было слабым, поскольку образование специфических комплексов рибозима с субстратом обнаружить не удалось.

Перенос ацильных групп, катализируемый РНК. Повсеместное использование аминоациладенилатов и аминоацилированных тРНК при трансляции в природных условиях указывает на возможную роль РНК в ранних системах трансляции. Действительно, способность рибозимов к спонтанному аминоацилированию позволяет предполагать и возможность их участия в следующем этапе биосинтеза белка – переносе аминокислотной группы с донорной молекулы на свою собственную акцепторную, т.е. в осуществлении аминокислоттрансферазных реакций. Такая активность была обнаружена среди 10^{15} молекул РНК случайной структуры, подвергнутых искусственному отбору на протяжении нескольких циклов. В этой системе отбора в качестве донора аминокислотной группы был использован гексануклеотидный фрагмент РНК, содержащий на 2'(3')-конце активированный остаток N-биотинилированного метионина. С помощью стрептавидин-агарозы отбирали молекулы РНК, способные переносить ацильную группу на свою собственную 5'-гидроксильную группу. Подобные рибозимы, среди которых доминировали однотипные молекулы, были получены после 11 циклов отбора.

Анализ первичной структуры таких рибозимов обнаружил вблизи 3'-концов наличие высококонсервативной внутренней матрицы, которая обладала способностью связывать и приводить в контакт друг с другом 2'(3')-конец донорной молекулы и акцепторный 5'-конец своей собственной полинуклеотидной цепи. Рибозимы оказались не просто матрицей. Они действительно *катализировали* перенос ацильных групп, ускоряя этот процесс в 10^3 раз ($k_{\text{кат}} = 9,4 \cdot 10^{-2} \text{ мин}^{-1}$) по сравнению с системами, в которых была задействована только матрица. Интересно, что два неправильно спаренных основания G-U в месте стыковки дуплексов донор-матрица и акцептор-матрица оказались необходимыми для функционирования рибозимов, стимулируя реакцию переноса ацильной группы в 10 раз.

Образование амидных связей, опосредованное рибозимами.

Рибозимы, конструирование которых было описано выше, катализировали реакцию трансэтерификации карбоксиэфиров. Для проверки возможности образования этими рибозимами амидных связей их 5'-концевые гидроксильные группы заменяли на аминогруппы. Оказалось, что такие рибозимы с высокой эффективностью осуществляют перенос биотинилированного метионина с 2'(3')-конца гексануклеотида-субстрата на свою собственную 5'-концевую аминогруппу. Константа скорости первого порядка реакции образования амидной связи ($0,58 \text{ мин}^{-1}$) была лишь в 15 раз ниже таковой (8 мин^{-1}), осуществляемой природной пептидилтрансферазой рибосом с использованием фрагмента аминокислотилированной тРНК.

РНК-центрический взгляд на происхождение трансляции и самой жизни. Получение аптамеров, способных специфически распознавать аминокислоты, и рибозимов, аминокислотилирующих самих себя, явилось серьезным аргументом в пользу ключевой роли РНК в происхождении трансляции. Если предположить, что РНК могут мобилизовать энергию макроэргических связей АТФ, то в скором времени будут обнаружены и рибозимы, активирующие аминокислоты путем синтеза соответствующих аминокислотаденилатов. Следуя тем же аргументам, можно надеяться на скорое получение рибозимов, способных синтезировать пептиды. Если все основные стадии синтеза белка будут осуществляться рибозимами, то, по мнению А. Хагер и соавторов (1996 г.), можно предположить следующие этапы эволюционирования системы трансляции. Вначале синтез пептидов

выполняется рибозимами, не обладающими высокой специфичностью по отношению к своим субстратам – донорам и акцепторам аминокислотных групп. Некоторые пептиды могли начать синтезироваться после приобретения рибозимами специфичности в отношении выбора субстратов. Хотя этот громоздкий механизм для образования каждой пептидной связи в синтезируемом пептиде требует собственного рибозима, однако он не кажется абсурдным, так как до сих пор бактерии используют такой принцип для синтеза некоторых своих пептидов (например молекулы грамицидина) с помощью специфического набора ферментов. Активность рибозимов могла быть повышена с помощью коротких РНК, объединяющих субстраты друг с другом. После эволюционного возникновения таких кофакторных РНК требования к специфичности рибозимов, синтезирующих пептиды, будут снижаться, и в конце концов появится истинная пептидилтрансфераза в виде 23S рРНК. Как известно, рРНК функционирует в тесной связи с матрицей мРНК, которая специфически сближает донорные и акцепторные аминокислот-тРНК.

Все рассмотренные аргументы подчеркивают важную, если не исключительную, роль РНК в происхождении жизни на Земле. Большинство современных ферментов являются белками, а все известные рибозимы действуют на РНК-субстраты. Уже одно это показывает, что цепь событий, приведшая к такому повороту развития эволюционных преобразований живых систем, остается до конца не понятой. Лабораторные методы, позволяющие моделировать эволюционирование рибозимов *in vitro*, дают возможность продолжить экспериментальное исследование глобального вопроса о происхождении жизни.

ГЛАВА 10. ТРАНСГЕННЫЕ ЖИВОТНЫЕ И РАСТЕНИЯ

Методические достижения генной инженерии микроорганизмов, разрушившие межвидовые барьеры в передаче генов, подготовили благоприятную почву для введения рекомбинантных ДНК в многоклеточные организмы. Действительно, идея введения чужеродных генов в геном многоклеточного организма ничем существенно не отличается от блестяще реализованных идей генетической трансформации бактериальных клеток гетерологичными последовательностями нуклеиновых кислот. Поэтому неудивительно, что с развитием генной инженерии появление трансгенных животных и растений не заставило себя долго ждать.

10.1. Способы получения трансгенных многоклеточных организмов

Многоклеточный организм высших животных и растений является продуктом онтогенетического развития, при котором из одной клетки (зиготы), образовавшейся в результате слияния двух половых клеток родителей (гамет), путем большого числа дроблений образуется вся совокупность высокодифференцированных клеток органов и тканей организма. Поскольку любая соматическая клетка или клетка зародышевого пути, в конечном счете, берет свое начало от двух объединившихся родительских клеток, она, как правило, заключает в себе всю (или большую часть) генетическую информацию родительских организмов. Несмотря на то что эта схема является упрощенной и по мере развития дифференцированного состояния соматических клеток их генетический материал часто претерпевает необратимые перестройки (например эритроциты человека вообще лишены ядер), такая картина подчеркивает преемственность генетического материала в рядах клеточных поколений соматических клеток организмов.

Почти все гены зигот имеют хорошие шансы быть представленными в большинстве соматических клеток организма и принять участие в формировании их генотипа и фенотипа. Предпосылки такого рода привели к мысли о возможности изменения фенотипа многоклеточных организмов путем введения новых рекомбинантных генов в геном зигот, еще не претерпевших дробления в раннем эмбриональном развитии. В случае объединения с

геномом зиготы новые гены должны распространиться в ряду клеточных поколений соматических клеток и экспрессироваться в большинстве этих клеток. Поскольку, с известными ограничениями, весь многоклеточный организм можно рассматривать как клон соматических клеток, произошедших от единственной клетки, распространение рекомбинантных генов, введенных в зиготу, в соматических клетках организма допустимо рассматривать как разновидность молекулярного клонирования последовательностей ДНК.

Такой молекулярно-генетический подход к изменению генотипа и фенотипа многоклеточных организмов был реализован экспериментально в середине 1970-х годов. Заражение мышинных эмбрионов на предимплантационной стадии развития вирусом лейкоза мышей (MuLV) приводило к образованию взрослых особей, содержащих вирусную ДНК, интегрированную в геном как соматических клеток, так и клеток зародышевого пути, и эта ДНК передавалась из поколения в поколение. Гены, искусственно введенные в геном многоклеточных организмов и передающиеся от родителей потомству, получили название *трансгенов*, процесс такого введения и передачи генов обозначили *трансгенозом*, а животные или растения, содержащие трансгены в геноме своих клеток, стали называть *трансгенными*. Развитие техники создания трансгенных животных и растений привело к возникновению нового быстро развивающегося направления молекулярной генетики. Были получены уникальные знания об особенностях экспрессии генов и биосинтезе белков в онтогенезе многоклеточных организмов, а также о возможности изменения фенотипа трансгенных организмов, в том числе и коррекции мутантного фенотипа, и использования трансгенных организмов для решения задач биотехнологии, связанных с биосинтезом рекомбинантных белков.

Для получения трансгенных животных в настоящее время применяют три основных метода. Во всех методах рекомбинантные гены в составе векторных молекул вводятся в клетки эмбрионов на ранних стадиях их развития. Наиболее популярным и эффективным методом является прямая микроинъекция нескольких сотен копий линеаризованных молекул рекомбинантной ДНК в пронуклеусы оплодотворенных яйцеклеток. Для этого мышей-самок скрещивают с самцами-производителями и через 12 ч после спаривания вскрывают их яйцеводы, выделяют оплодотворенные

одноклеточные яйца и помещают их в культуральную жидкость. В пронуклеусы оплодотворенных яиц с помощью микроманипулятора вводят очищенную ДНК, и прооперированные яйца пересаживают в яйцеводы псевдобеременных реципиентных самок. Некоторая часть жизнеспособных трансплантированных яиц проходит в организме мышей полный пренатальный цикл развития, после чего развившиеся детеныши рождаются естественным путем или с использованием кесарева сечения. Трансгенных животных идентифицируют в помете гибридизацией по Саузерну высокомолекулярной ДНК тканей хвостов с соответствующими зондами, далее трансгенные животные скрещивают друг с другом для получения чистых линий и анализируют экспрессию трансгенов. Приблизительно у 70% трансгенных мышей экзогенная рекомбинантная ДНК имеется во всех соматических клетках и клетках зародышевого пути. Это свидетельствует о том, что у них интеграция рекомбинантной ДНК в хромосомы прошла до прохождения первого цикла их репликации в оплодотворенной яйцеклетке. У остальных 30% трансгены содержат лишь часть соматических клеток, т.е. они являются мозаиками, причем у некоторых трансгенных животных клетки зародышевого пути становятся дефектными, а сами животные – бесплодными. Как правило, трансгены стабильно передаются из поколения в поколение без существенных изменений. Однако в ряде случаев отмечены их перестройки, образование делеций и амплификация. Кроме непосредственной микроинъекции в пронуклеусы используют введение рекомбинантной ДНК в цитоплазму или ядра двухклеточных эмбрионов, а также в полость бластоцеля зародышей.

Прямая микроинъекция рекомбинантных генов в клетки высокоэффективна, однако методически сложна и требует дорогостоящего оборудования. Более простым способом доставки чужеродных генов в геном животного-реципиента является использование векторов на основе вирусов. В этом случае эмбрионы на ранней (восьмиклеточной) стадии развития инкубируют в культуральной среде в присутствии фибробластов, в которых образуются рекомбинантные ретровирусы, и после заражения такими вирусами эмбрионы пересаживают псевдобеременным самкам мышей, где они продолжают свое развитие. Кроме простоты одним из преимуществ данного способа введения ДНК является то, что в геном клеток зародышей интегрируется, как правило, одна копия исследуемого гена, фланкированного

длинными концевыми повторами вирусной хромосомы, что может способствовать эффективной экспрессии гена. Однако к недостаткам метода следует отнести необходимость проведения дополнительных генно-инженерных манипуляций при подготовке ретровирусного вектора, ограниченную емкость вектора (размер вставки – до 10 т.п.о.) и мозаицизм образующихся трансгенных животных, которые состоят из клеток как содержащих, так и не содержащих трансгены.

Наконец, следует упомянуть еще об одном способе введения рекомбинантных генов в клетки зародышевой линии животных, при котором используют плюрипотентные (дифференцирующиеся по разным направлениям) эмбриональные стволовые клетки линий ES и EK, предшественники которых были взяты из бластоцисты зародышей мышей. Рекомбинантные гены вводят в такие клетки любым из вышеупомянутых способов, а кроме того, электропорацией или другими стандартными методами, применяемыми для доставки генов в культивируемые соматические клетки. При этом вместе с исследуемыми генами возможно введение селективируемых маркеров, которые позволяют проводить отбор клеток, экспрессирующих данные маркеры и, следовательно, гарантированно содержащих сцепленные с ними исследуемые гены. Отобранные таким образом клетки переносят в бластоцисты развивающихся эмбрионов или используют для получения агрегационных химер объединением их с клетками восьмиклеточных эмбрионов с последующей пересадкой эмбрионов псевдобеременным самкам.

10.2. Экспрессия трансгенов

Если трансгены в своем функционировании проявляют тканеспецифичность, то уровень их экспрессии зависит от места интеграции в хромосому. В тех редких случаях, когда экспрессия трансгена полностью отсутствует, это объясняют его интеграцией в гетерохроматиновые участки хромосом, неактивные в отношении транскрипции, или другими эффектами положения. Поскольку уровень экспрессии трансгенов у некоторых трансгенных мышей может превышать таковой у его эндогенного гомолога, делают вывод об отсутствии необходимости в точной локализации гена на хромосоме для его эффективной экспрессии. Тканеспецифический характер экспрессии генов

обеспечивают, главным образом, их энхансеры – *цис*-действующие регуляторные последовательности нуклеотидов, которые располагаются как внутри, так и вне генов на значительном удалении от них (см. рис. 1.30). Обнаружены гены, экспрессирующиеся в клетках только одного типа (строгая тканеспецифическая экспрессия), в клетках немногих тканей и в клетках многих или большинства типов (так называемые гены "домашнего хозяйства"). В том случае, если энхансеры обеспечивают абсолютный тканеспецифический характер экспрессии, контекст окружающих трансген последовательностей нуклеотидов хромосомы оказывает влияние на уровень его экспрессии только в клетках одного типа. Если же энхансеры дают возможность функционировать трансгену в клетках разных тканей, то уровень экспрессии будет варьировать в клетках этих тканей даже при интеграции в геном единственной его копии.

Обнаружена слабая корреляция между числом копий трансгенов, тандемно интегрированных в геном животного, и суммарным уровнем их экспрессии. Это указывает на функциональную активность лишь немногих трансгенов из всего кластера их множественных копий.

Таким образом, данные, полученные к настоящему времени, указывают на то, что экспрессия рекомбинантных генов у трансгенных животных в большей степени напоминает таковую, происходящую в природных условиях в гомологичном генетическом окружении. Объединение структурных частей рекомбинантных генов с конкретными регуляторными последовательностями организма-реципиента позволяет получать строго детерминированную, тканеспецифическую экспрессию этих генов в процессе трансгеноза.

10.3. Использование трансгенов у животных

Техника трансгеноза открывает практически безграничные, принципиально новые возможности исследования экспрессии генов. Ниже будут кратко рассмотрены четыре активно развиваемые направления использования трансгенов в фундаментальных и прикладных исследованиях. К ним относятся: а) изучение молекулярных механизмов экспрессии генов; б) исследование механизмов дифференцировки соматических клеток в онтогенезе многоклеточных организмов; в) подходы к изменению физиологического статуса лабораторных и сельскохозяйственных животных; г) моделирование наследственных и приобретенных заболеваний человека. Два

последних направления исследований являются фундаментом для современной и будущей генотерапии.

10.3.1. Исследование механизмов экспрессии генов

Как уже упоминалось выше, *цис*-действующие регуляторные последовательности нуклеотидов обеспечивают тканеспецифический характер экспрессии трансгенов. Этим свойством воспользовались для определения точной локализации таких регуляторных элементов генов. Например, с использованием делеционного картирования был идентифицирован энхансерный элемент гена эластазы-1 крыс. Полный ген эластазы-1 в составе фрагмента ДНК длиной в 7,2 т.п.о. экспрессировался у трансгенных мышей, что указывало на наличие в данном фрагменте ДНК всех последовательностей нуклеотидов, необходимых для функционирования гена. Удаление с помощью делеций 5'-фланкирующих последовательностей нуклеотидов этого гена не влияло на тканеспецифический характер его экспрессии в ацинарных клетках поджелудочной железы мышей (место обычной активации гена) до тех пор, пока не затрагивалась последовательность длиной в 205 п.о., непосредственно прилегающая к 5'-концевой его части (рядом с точкой инициации транскрипции). Делеции, которые оставляли лишь 72 п.о. 5'-фланкирующей последовательности, полностью инактивировали ген, что указывало на расположение регуляторного элемента между нуклеотидами в положениях -205 и -72. Такая последовательность обладала всеми свойствами энхансера и сохраняла активность после перемещения за 3 т.п.о. перед сайтом ее нормальной локализации или встраивания в интрон. При этом для нее была характерна способность активировать гетерологичные промоторы.

Использование таких эффективных, хотя и трудоемких методов иногда является единственным путем идентификации тканеспецифических регуляторных элементов генов. Дело в том, что клетки в культуре тканей часто претерпевают дедифференцировку и утрачивают изначально присущий им тканеспецифический характер экспрессии своих генов, поэтому применение культур тканей вместо трансгенных животных для таких целей проблематично.

Экспрессия трансгенов обнаруживается по появлению соответствующего им белкового продукта или транскрипта, а также по специфическим физиологическим реакциям организмов-реципиентов. Это позволяет

использовать трансгеноз для изучения путей и механизмов дифференцировки соматических клеток в биологии развития, а также для изменения физиологической и физической конституции сельскохозяйственных животных, моделирования многих наследственных заболеваний человека и разработки методов их лечения.

10.3.2. Токсигены в исследовании дифференцировки соматических клеток в онтогенезе

Выше рассматривалась возможность применения гибридных токсинов для специфического воздействия на группы соматических клеток, обладающих определенными фенотипическими маркерами. Высокоспецифической элиминации определенных групп дифференцирующихся клеток можно достигнуть с использованием трансгеноза. В этом случае гены каталитических субъединиц токсинов (например А-субъединиц дифтерийного токсина (DT-A) или рицина) с помощью генно-инженерных приемов помещают под контроль регуляторных генетических элементов, обеспечивающих их тканеспецифическую экспрессию в клетках трансгенных экспериментальных животных.

После интеграции таких *токсигенов* в геном развивающегося эмбриона они передаются от клетки к клетке в рядах клеточных поколений без экспрессии до тех пор, пока клетки не начинают дифференцироваться с образованием регуляторных белков, трансктивирующих регуляторные последовательности токсигенов. С этого момента токсигены активированы, т.е. приобрели способность направлять биосинтез белков-токсинов, которые избирательно поражают дифференцирующиеся клетки. Иными словами, начав дифференцироваться в определенном направлении, соматические клетки совершают самоубийство. Поскольку каталитические субъединицы обладают токсическим эффектом лишь внутри соматических клеток и сами по себе не могут проникнуть в окружающие их клетки извне, их токсический эффект распространяется только на клетки, в которых происходит экспрессия токсигенов.

С помощью такой стратегии были получены впечатляющие результаты при изучении путей терминальной дифференцировки соматических клеток животных *in vivo*. В частности, были продемонстрированы различия во

временной активации промоторов генов $\gamma 2$ - и $\alpha 2$ -кристаллинов в процессе формирования хрусталиков глаз у мышей, а также последствия активации промоторов гена эластазы-I, специфически функционирующего в ацинарных клетках поджелудочной железы.

Из-за большой токсичности вышеупомянутых белковых токсинов (чтобы убить клетку, достаточно одной молекулы) использование таких токсигенов может сопровождаться артефактами и приводить к ранней гибели трансгенных эмбрионов. Для преодоления этих трудностей были разработаны методы, при которых сам по себе трансген не оказывает токсического действия на клетки, но внутриклеточное присутствие белкового продукта его экспрессии делает клетки чувствительными к определенному экзогенному воздействию.

В одной из таких систем использовали ген тимидинкиназы вируса простого герпеса (HSV-tk), продукт которого, в отличие от тимидинкиназы клеток-хозяев, обладает способностью фосфорилировать некоторые аналоги нуклеозидов, например ацикловир и FIAU [1-(2-дезоксидеокси-2-фтор- β -D-арабинофуранозил)-5-иодоурацил]. После фосфорилирования с помощью HSV-tk до нуклеозидмонофосфатов образующиеся аналоги нуклеотидов могут далее фосфорилироваться тимидилаткиназой клеток-хозяев с образованием соответствующих дезоксирибонуклеозидтрифосфатов и конкурентно блокировать репликацию ДНК соматических клеток. Аналогичный подход был использован для элиминации в эмбриогенезе клеток, экспрессирующих кристаллины, кроветворных клеток (в частности тимоцитов) и в ряде других случаев.

Токсигены находят применение не только для изучения генетики развития, но и для создания животных, моделирующих наследственные и приобретенные заболевания человека.

10.3.3. Изменение физиологического статуса лабораторных и сельскохозяйственных животных

Одними из первых указаний на возможность использования трансгеноза для изменения физиологических параметров и физической конституции организма животных были результаты работ по экспрессии трансгенов гормонов, контролирующих рост млекопитающих: рилизинг-фактора гормона

роста (GRF), соматостатина и самого гормона роста (growth hormone, GH). Гормон роста индуцирует образование в печени инсулиноподобного фактора роста I (IGF-I), который, в свою очередь, ускоряет пролиферацию клеток. Повышенное внутриклеточное содержание GH крысы, быка или человека в клетках трансгенных мышей сопровождалось ускорением их роста через три недели после рождения, рост стабилизировался через 12 недель, когда размер мышей в два раза превышал обычный. К аналогичному эффекту приводила экспрессия у трансгенных мышей гена GRF человека, находящегося под контролем промотора гена металлотионеина, который стимулировал эндогенное образование GH, что сопровождалось повышением внутриклеточного уровня IGF-I-mPHK и ускорением роста трансгенных животных.

Увеличение уровня экспрессии GH у трансгенных мышей обуславливает и ряд других физиологических реакций организма. Например, под действием трансгена GH отмечены ингибирование биосинтеза эндогенного GH и дегенерация соматотропных клеток гипофиза, в норме продуцирующих этот гормон. Экспрессия приводила, кроме того, к изменению половых различий в строении печени, ослаблению фертильности самок и преждевременному старению трансгенных животных.

Результаты таких опытов легли в основу целого направления прикладных исследований, в которых были предприняты попытки изменения физиологического статуса сельскохозяйственных животных путем гормональных воздействий, опосредованных трансгенами. Были получены трансгенные овцы и свиньи, однако результаты оказались не вполне удовлетворительными из-за вышеупомянутых побочных эффектов GH. Решение этой проблемы, по-видимому, лежит на пути создания трансгенных животных с регулируемой экспрессией трансгенов. М. Мак-Грейну с соавторами недавно удалось получить гигантских мышей с помощью трансгена GH быков, экспрессирующегося под контролем регуляторной области крысиного гена 3-энолпируваткарбоксикиназы, для которого характерна тканеспецифическая экспрессия в печени и почках. Авторам удалось управлять экспрессией трансгена путем изменения содержания углеводов и белков в пищевом рационе мышей. Такой же подход был успешно применен для получения полноценных трансгенных свиней, которые характеризовались ускоренным

увеличением веса и существенным уменьшением содержания жира в тканях.

Другое направление биотехнологии, в котором используется трансгеноз, – изучение возможности биосинтеза чужеродных белков в клетках молочных желез млекопитающих с последующей их секрецией в молоко. Идея этого подхода заключается в применении молочных желез крупных сельскохозяйственных животных в качестве биореакторов для биосинтеза биологически активных пептидов и белков. Предполагалось, что в результате экспрессии трансгенов в клетках молочных желез удастся синтезировать рекомбинантные белки, претерпевшие правильные посттрансляционные модификации (в частности гликозилирование), что, как правило, не удастся делать в других генетических системах.

Уже в первых опытах с трансгенами тканевого активатора плазминогена (tPA) человека, интегрированными в геном в виде кДНК, была показана возможность специфической экспрессии его в молочных железах мышей, причем содержание функционально активного tPA составляло около 0,1 мг на 1 мл. Похожие результаты были получены с трансгенами фактора IX коагуляции крови, α 1-антитрипсина, γ -интерферона, фолликулостимулирующего гормона и интерлейкина 2 человека, а также ряда других белков, имеющих значение для биотехнологии.

Во всех вышеперечисленных направлениях исследований усилия были сосредоточены на получении трансгенных животных-биореакторов, секретирующих в молоко белковые продукты экспрессии трансгенов, что предполагает их дальнейшую очистку. В отличие от этого в других исследованиях предпринимались попытки изменения формулы и питательных свойств молока. В частности, Дж. Симмонсу с соавторами удалось получить экспрессию гена β -лактальбумина овец в молочных железах трансгенных мышей. Авторы не исключали принципиальной возможности сделать состав молока крупных сельскохозяйственных животных более пригодным для новорожденных детей или людей, обладающих повышенной чувствительностью к лактозе.

10.3.4. Моделирование наследственных и приобретенных заболеваний человека

Для разработки эффективных методов лечения наследственных и приобретенных заболеваний человека, а также для полного понимания их этиологии требуется моделирование соответствующих симптомов на лабораторных животных. В этом случае проблема заключается в направленном введении мутаций в гены организма животных, инактивация которых приводит к развитию патологических процессов. Разработка эффективных методов получения трансгенных животных позволила исследователям вплотную подойти к решению данного вопроса.

Основой для решения послужили две работы, выполненные М. Хупером и М. Куэном с соавторами и опубликованные в 1987 г. Авторам удалось разработать общий подход к избирательной инактивации генов в организме животных. Мишенью для инактивации стал ген гипоксантин-гуанозин-фосфорибозилтрансферазы (ГГФРТ), нефункциональное состояние которого у человека приводит к развитию заболевания Леша–Нихана. Мышей, дефектных по гену ГГФРТ, получали из эмбриональных стволовых клеток линии ES, инактивируя ген с помощью спонтанных мутаций или интегрируя в него геном ретровирусов (разновидность инсерционного мутагенеза). Клетки с инактивированным геном ГГФРТ удобно отделять от клеток дикого типа на селективной питательной среде в присутствии 6-тиогуанина. Позднее было показано, что инактивация гена ГГФРТ может быть достигнута путем гомологичной рекомбинации гена дикого типа с мутантным геном или его частью, которые вводят в клетки ES с помощью электропорации или микроинъекций в составе линеаризованных векторных плазмид. Основная проблема, возникающая при инактивации гена-мишени с помощью гомологичной рекомбинации, заключается в низкой частоте ($\sim 10^{-6}$) рекомбинационных событий, приводящих к правильной интеграции экзогенной последовательности нуклеотидов в инактивируемый ген. Эта проблема в настоящее время решается путем отбора требуемых мутантных клеток из общей популяции клеток, в которых интеграция инактивирующего вектора в хромосому произошла случайным образом.

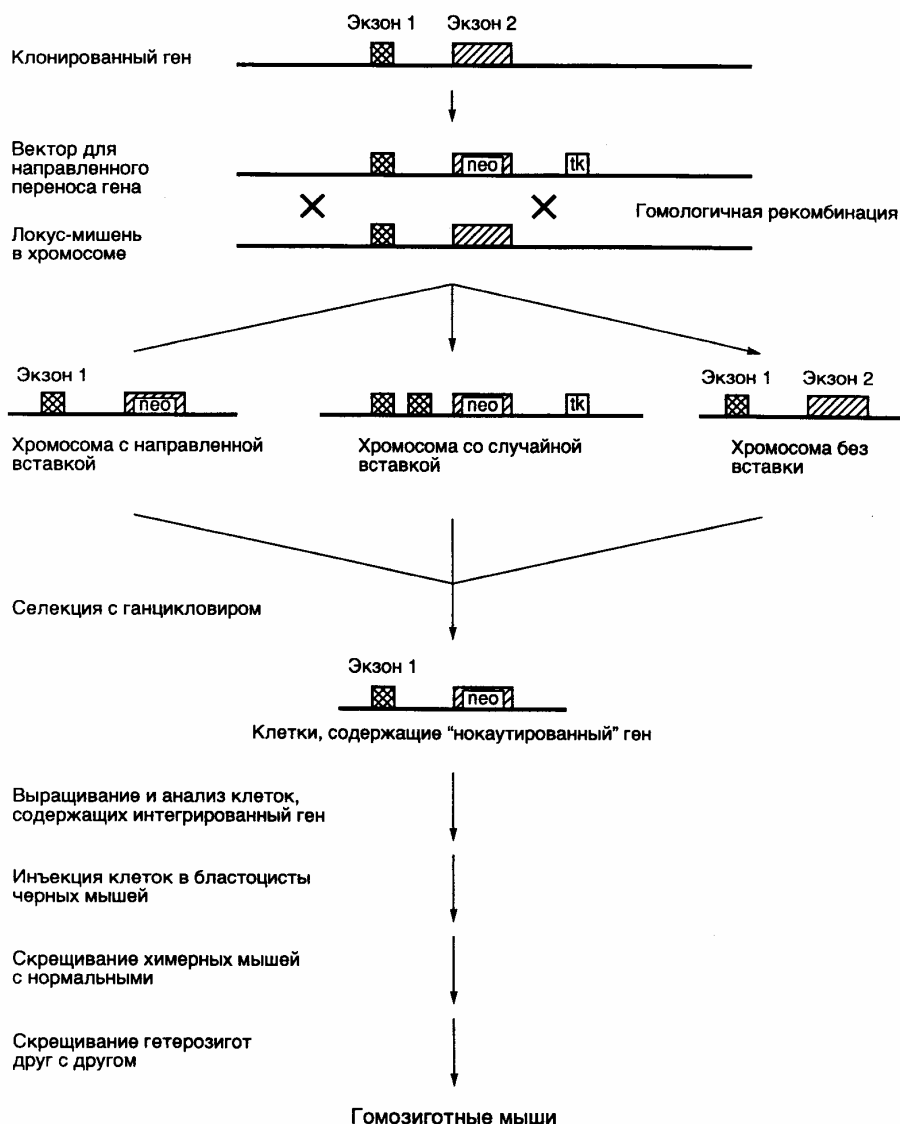


Рис. II.29. Схема адресной доставки генов в геном эмбриональных стволовых клеток с использованием гомологичной рекомбинации ("генный нокаут")

Показано положение гена устойчивости к неомизину (neo), интегрированного в экзон 2 инактивируемого гена, который используется для позитивного отбора, а также гена тимидинкиназы вируса простого герпеса (tk), используемого для негативной селекции клеток в присутствии ганцикловира

Чтобы создать селективные условия отбора, в последовательности нуклеотидов рекомбинантной ДНК, гомологичные последовательностям инактивируемого гена, вводят доминантный селектируемый маркер в виде гена

устойчивости к антибиотикам (например *leo*), который экспрессируется только в случае правильной интеграции в ген-мишень под контроль его регуляторных элементов. Другая, более эффективная система отбора основана на одновременном проведении негативной и позитивной селекции клеток, у которых произошла правильная интеграция инактивирующего вектора. В этой системе помимо доминантного селектируемого маркера *leo* инактивирующий вектор содержит другой маркерный ген, экспрессия которого приводит к гибели клеток ES на селективной среде. Маркерным геном часто является ген тимидинкиназы вируса простого герпеса (HSV-tk), принцип действия которого был рассмотрен выше. Ген HSV-tk вводится в инактивирующий вектор таким образом, что он удаляется в результате гомологичной рекомбинации и сохраняется в хромосоме в случае неспецифической интеграции рекомбинантной ДНК, что приводит к цитотоксическому эффекту. Обобщенная схема направленного введения мутаций в гены животных *in vivo*, которые приводят к их инаktivации во всех клетках трансгенного организма ("*генный нокаут*"), представлена на рис. II.29.

Разработка универсальных методов направленной инаktivации любого требуемого гена во всех клетках организма трансгенных животных позволила за короткое время создать модели таких наследственных заболеваний человека, как β -талассемия, мышечная дистрофия Дюшенна, серповидно-клеточная анемия, муковисцидоз, болезнь Леша–Нихана и целого ряда других распространенных синдромов. Создание этих методов открыло также возможности генотерапии наследственных заболеваний путем замещения в клетках зародышевой линии мутантных аллелей на нормальные. Результаты, полученные в этой области, будут обсуждаться в разделе 10.5.

Следует подчеркнуть, что инаktivация генов-мишеней у лабораторных животных – не единственный способ моделирования наследственных и особенно приобретенных заболеваний человека, поскольку причиной многих заболеваний является не отсутствие функционирования, а сверхэкспрессия определенных генов. Например, высокий уровень экспрессии трансгена аполипопротеина All сопровождается развитием острого атеросклероза у трансгенных мышей. Только исчерпывающее знание этиологии заболевания допускает его адекватное моделирование, и, наоборот, лишь создание адекватных моделей может убедить исследователя в том, что он правильно

понимает его природу.

10.4. Трансгенные растения

Способность к вегетативному размножению отличает организм растений от организма высших животных, что заметно облегчает осуществление трансгеноза. Многие клетки растений, например клетки зародыша на ранних стадиях его развития, покоящиеся клетки меристем кончиков побегов и корней, а также сосудистых тканей камбия, находятся в недетерминированном состоянии и, попадая под влияние внешних воздействий, могут дифференцироваться с образованием клеток любых типов, а также давать начало новым растениям. В частности, перенос в питательную среду таких недетерминированных клеток может приводить к их полной дедифференцировке и формированию в культуре недифференцированной ткани каллуса. Такие клетки могут стабилизироваться в жидких суспензионных культурах и расти неограниченно долго. Из недифференцированных тканей многих видов растений можно легко регенерировать целые растения.

Процесс получения трансгенных растений в этом случае начинается с введения требуемых генов в недифференцированные клетки таким образом, чтобы они интегрировались в хромосомы. Введение чужеродных генов в клетки растений облегчается, если их клеточные стенки удаляют с помощью гидролитических ферментов - пектиназы и(или) целлюлазы, что приводит к образованию протопластов. Чужеродные гены, находящиеся в составе векторных плазмид, вводят в протопласты одним из стандартных способов с использованием эндоцитоза, стимулированного полиэтиленгликолем, электропорации, микроинъекций или бомбардировки микрочастицами, нагруженными векторной ДНК. После этого протопласты в течение нескольких дней культивируют на питательной среде для восстановления клеточных стенок и образующиеся клетки-трансфектанты используют для регенерации целых растений.

Основным направлением применения трансгеноза для генетической модификации культурных растений является повышение их устойчивости к неблагоприятным воздействиям окружающей среды, в частности вирусам и гербицидам. Один из таких подходов, в котором антисмысловые РНК были направлены на подавление экспрессии жизненно важных генов вирусов, уже

был рассмотрен в разделе 9.1.2.

Другой метод защиты растений от вирусов с помощью трансгенов, предложенный В. Шибальским в 1988 г., успешно используется в настоящее время. Сущность метода заключается во введении в геном растений *транс-действующих* доминантных летальных генов или, по терминологии Шибальского, "*анти-генов*", которые кодируют измененные мутациями белки вирусов, существенные для их воспроизводства, и путем конкурентного замещения соответствующих белков вируса дикого типа прерывают его размножение. В частности, с использованием такого подхода удалось получить очень высокую устойчивость растений к вирусу Х картофеля (PVX). В этом случае в ген репликазы PVX с помощью направленного мутагенеза вводили мутации, сопровождающиеся заменой аминокислот в консервативном участке полипептидной цепи репликазы, ассоциированном с ее каталитическим сайтом. Для экспрессии мутантного трансгена в растениях табака были характерны внутриклеточное накопление инактивированной репликазы и появление высокой устойчивости растений к заражению вирусом PVX.

Еще один современный подход к получению трансгенных растений, устойчивых к вирусам, основан на введении в них трансгенов, экспрессирующих в клетках моноклональные антитела, направленные против вирусных белков. В одной из работ с использованием такого метода создали эффективную систему защиты растений от вируса морщинистой мозаики артишока (AMCV). Для этого сначала получили панель моноклональных антител к вирусу AMCV и отобрали гибридомы, продуцирующие антитела, которые взаимодействуют с консервативными участками белка оболочки вируса. Клетки гибридомы использовали для конструирования библиотеки кДНК, из которой выделили последовательности нуклеотидов, кодирующие полноразмерные тяжелые и легкие цепи иммуноглобулинов G класса 2b. С помощью ПЦР и универсальных праймеров амплифицировали вариабельные участки этих последовательностей (V_H и V_L), которые далее клонировали в экспрессирующий вектор *E. coli*, что сопровождалось образованием полипептидов V_H и V_L , соединенных линкерным пептидом (антитела scFV). После отбора клонов, продуцирующих высокоаффинные антитела к вирусному антигену (scFV), объединенные таким образом гены V_H - V_L помещали в экспрессирующий вектор и использовали для получения трансгенных растений

табака *Nicotiana benthamiana*. Трансгенные растения содержали в своих клетках до 0,1% антител от суммарного белка и оказались устойчивыми к AMCV-инфекции, но не к вирусу мозаики цветной капусты (CMV), что указывало на специфический характер их резистентности.

В заключение следует упомянуть о работе, в которой трансгенные растения сорго, устойчивые к гербицидам, получали бомбардировкой незрелых эмбрионов на стадии зиготы микрочастицами золота (диаметр частиц – 1,5–3,0 мкм). В таком случае микрочастицы погружали в раствор экспрессирующего вектора, высушивали и "выстреливали" в клетки-мишени, добиваясь при этом высоких результатов трансфекции.

10.5. Генотерапия наследственных и приобретенных заболеваний

Современные методы лечения наследственных и приобретенных заболеваний связаны с введением в организм больного недостающих продуктов метаболизма или с ограничением поступления их предшественников с пищей, как это происходит, например, в случае диабета и т.п. Практикуется введение в организм и самих продуктов экспрессии поврежденного гена, а именно: ферментов или пептидных гормонов. Естественно, что подобная терапия лишь временно снимает симптомы заболевания и не устраняет его причины. Радикальным способом лечения таких заболеваний было бы исправление генетического дефекта, приводящего к развитию заболевания, путем введения в геном организма функционально активного гена. Современный уровень развития генной инженерии в принципе позволяет разрабатывать подходы к решению возникающих в связи с этим проблем, поскольку основные методы получения трансгенных животных и растений могут быть использованы и для манипулирования генами человека.

Теоретически возможны два типа генно-инженерных воздействий на геном человека с целью генотерапии. Во-первых, это вмешательство на уровне соматических клеток, приводящее к созданию клона клеток с измененным генотипом, которые могут повлиять на фенотип организма в целом. Во-вторых, воздействие на клетки зародышевой линии. В этом случае все клетки организма следующего поколения будут генетически изменены и такие изменения будут передаваться из поколения в поколение. Объектами генно-инженерных воздействий служат как половые клетки до оплодотворения, так и

эмбриональные клетки на ранних стадиях развития зародыша до его имплантации. При работе с соматическими клетками достаточно введения дополнительной копии нормального гена, который может существовать в геноме наряду с его поврежденным гомологом. В отличие от этого в случае генотерапии на уровне клеток зародышевой линии необходимо производить замещение мутантного гена нормальным, так как современный уровень знаний не позволяет предсказать фенотипические последствия одновременного присутствия в геноме нормального и мутантного аллелей гена и их влияние на процессы роста и развития организма.

В настоящее время генотерапия рассматривается как многообещающее направление в развитии способов лечения широкого круга заболеваний человека как моногенной, так и полигенной (многофакторной) этиологии. При этом обработку клеток-мишеней трансгенами можно осуществлять как в организме человека, так и *ex vivo*, т.е. за его пределами. В последнем случае соматические клетки (чаще всего крови) больного подвергают генно-инженерному воздействию вне организма, после чего вводят их обратно. При всех этих генно-инженерных манипуляциях с трансгенами в первую очередь обращают внимание на генетическую безопасность подобных операций т.е. предотвращение нежелательных побочных эффектов трансгенов. Кроме того, после направленной доставки трансгенов к клеткам-реципиентам необходимо обеспечивать контроль за уровнем их экспрессии, а также, в идеальном случае, возможность изменения этого уровня, т.е. регулировать экспрессию трансгенов. Рассмотрим подробнее современную технику введения трансгенов при генотерапии, чаще всего используемую в клиниках, а также методы, обеспечивающие контроль их экспрессии в организме больного.

10.5.1. Способы доставки новых генов в геном человека

Ретровирусные векторы. Для доставки трансгенов в организм человека в целях генотерапии ретровирусные векторы используются наиболее широко и являются одним из наиболее эффективных средств доставки генетического материала в геном человека. Специфичность взаимодействия ретровирусов с поверхностью клеток-мишеней в момент заражения обеспечивается белком оболочки их вириона, кодируемого геном *env*, продукт которого контактирует с белком-рецептором заражаемых клеток. Это создает предпосылки для

направленного изменения круга хозяев вируса, т.е. клеток, которые он может заражать, путем изменений белка вирусной оболочки. Многие ретровирусы обладают ограниченным кругом хозяев вследствие функционирования вышеупомянутого механизма. Классическим примером такого рода является вирус иммунодефицита человека (ВИЧ), который способен заражать лишь субпопуляцию лимфоцитов, экспрессирующих на своей поверхности рецептор CD4.

Наиболее пристальное внимание как к средству доставки трансгенов при генотерапии уделяется вирусу лейкоза мышей Молони (MoMLV). Вирусы группы MoMLV-E обладают способностью заражать практически все исследованные клетки грызунов (вирусы с кругом хозяев, который не выходит за рамки организмов, обычно заражаемых такими вирусами, получили название *экотропных*). В отличие от этого вирусы группы MoMLV-A способны заражать большое число клеток млекопитающих, включая клетки человека (вирусы с широким кругом хозяев получили название *амфотропных*). В настоящее время разрабатываются три основные стратегии искусственного изменения тропизма (круга хозяев) вирусов этих групп, что необходимо для доставки трансгенов в клетки человека: прямое изменение последовательности нуклеотидов гена белка оболочки ретровирусов, соединение белка оболочки с новыми лигандами и псевдотипирование.

Модификации гена белка оболочки. Последовательности нуклеотидов гена *env*, отвечающие за взаимодействие с клеточными рецепторами, определены и могут быть непосредственно замещены последовательностями, кодирующими лиганды невирусной природы. С использованием этого подхода удалось получить химерные белки оболочки, содержащие последовательности аминокислот, изменяющие тропизм вируса. В частности, слияние эпидермального фактора роста (ЭФР) с белком оболочки амфотропного ретровируса через расщепляемую протеиназой фактора Ха линкерную последовательность аминокислот приводило к взаимодействию вирусов преимущественно с ближайшими клетками, содержащими на своей поверхности рецепторы фактора. После отщепления ЭФР протеиназой вирусные частицы начинали заражать клетки, экспрессирующие гомологичные вирусные рецепторы. Однако необходимо иметь в виду, что специфическое взаимодействие ретровирусов с рецепторами на поверхности клеток-хозяев

требует образования контактов между несколькими частями полипептидной цепи белка оболочки и рецептора. Кроме того, белок оболочки ретровирусов не только обеспечивает контакт с рецепторами на поверхности клеток-мишеней, но и участвует в последующих этапах проникновения вируса внутрь клеток – интернализации в составе комплекса с рецептором. Все это затрудняет практическое использование подхода с использованием прямого замещения последовательностей аминокислот белка оболочки новыми последовательностями для эффективного изменения тропизма ретровирусов.

Соединение белка оболочки с новыми лигандами. Первые успехи на этом пути были достигнуты путем создания конъюгатов антител с белком оболочки ретровирусов с последующим соединением таких антител с антителами, специфичными в отношении антигенов поверхности клеток посредством стрептавидина. При данном подходе изменение тропизма ретровирусов было достигнуто при использовании антител к антигенам главного комплекса гистосовместимости, ЭФР или рецепторам трансферрина, экспрессирующимся на поверхности клеток гепатомы. Однако после интернализации вирусных частиц интеграция вирусного генома в геном зараженных клеток проходила крайне неэффективно. Предполагают, что рецептор в комплексе со связанным с ним агентом после проникновения в клеточный компартмент блокирует доступ вирусного генома к клеточному ядру, поскольку при обычной инфекции он задерживается и деградирует на поверхности клеток.

Второй подход в этой группе методов использует химическое присоединение лактозы к белкам оболочки или *десиамирование* гликопротеинов оболочки вируса. Оба типа модификации дают возможность модифицированным вирусным частицам специфически взаимодействовать с асиалогликопротеиновыми рецепторами, присутствующими, например на поверхности гепатоцитов. Модифицированные таким способом экотропные вирусы приобретают способность с высокой эффективностью заражать гепатоциты человека.

Псевдотипирование. При использовании этой группы методов упаковка геномной РНК ретровирусов проходит в культивируемых клетках, которые экспрессируют гетерологичный белок оболочки, замещающий обычный вирусный белок в процессе внутриклеточного формирования ретровирусных

частиц. Наиболее легко замещение белка вирусной оболочки происходит при использовании гомологичных белков близкородственных вирусов. Например, геном MoMLV может быть упакован в вирусные частицы с участием белков оболочки ретровирусов C-, но не D-типа. Неэффективным в этом процессе оказывается и белок оболочки ретровируса HTLV-1. В последнем случае он включается в оболочку только в присутствии всех остальных белков дикого типа вируса MoMLV. В практических целях широкое распространение получили клеточные линии, в которых геном рекомбинантного ретровируса упаковывается в оболочку амфотропного ретровируса (с широким кругом хозяев), например MoMLV-A. В другой популярной линии клеток происходит экспрессия белка оболочки вируса лейкоза гиббонов, а также гомологичных генов *gag-pol* вируса MoMLV. Подобное сочетание вирусных белков позволяет получать рекомбинантный вирус, пригодный для генотерапии, в высоком титре.

Эффективный перенос ретровирусных векторов происходит только в активно делящиеся клетки, экспрессирующие на своей поверхности соответствующие рецепторы. Стабильная интеграция ретровирусных векторов в геномную ДНК клеток-хозяев создает условия для длительной и эффективной экспрессии рекомбинантных генов, замещающих поврежденные аналоги. Вероятность того, что интеграция вектора в геном приведет к активации какого-либо онкогена, мала, и до сих пор это не наблюдалось в эксперименте. Размножение таких вирусов внутри клеток исключается самой процедурой их конструирования, так как они дефицитны по репликации. Ретровирусные векторы пригодны как для терапии *ex vivo*, так и для прямого переноса рекомбинантных генов в клетки реципиентов *in vivo*.

Аденовирусные векторы. В отличие от ретровирусов аденовирусы, в капсид которых упакованы рекомбинантные ДНК, способны заражать и неделящиеся клетки. Кроме того, этот тип векторов не интегрируется в геном клеток-реципиентов, в связи с чем их экспрессия внутри клеток носит временный характер. Векторы на основе аденовирусов используются в генотерапии реже.

Капсид аденовирусов представляет собой икосаэдр (правильный 20-гранник), каждая грань которого (гексон) составлена из шести идентичных белковых субъединиц, с его вершинами соединены пентоны (пентасубъединичные белки), к которым нековалентно N-концевой частью

присоединены гомотримерные стержни (knobs) с глобулярными доменами на С-концах. Пять идентичных субъединиц, составляющих пентоны, содержат по одному мотиву из трех аминокислот Arg–Gly–Asp (RGD – в однобуквенном обозначении), специфически взаимодействующих с рецепторами (*интегринами*). Это дает возможность аденовирусам осуществлять контакт с интегринными $\alpha_v\beta_3$ и $\alpha_v\beta_5$ на поверхности клеток. Взаимодействие аденовирусов с клетками-мишенями происходит в два этапа. Вначале глобулярные домены стержней связываются с первичными (пока не охарактеризованными) рецепторами на поверхности клеток-мишеней, затем после взаимодействия пентонов с интегриновыми рецепторами происходит интернализация вирусных частиц. Одна из проблем специфичности доставки трансгенов с помощью аденовирусов заключается в том, что многие клетки обладают вышеупомянутыми первичными рецепторами для глобулярных доменов стержней аденовирусов. В связи с этим проводятся работы по изменению специфичности связывания глобулярных доменов стержней генно-инженерными методами. При таком подходе делаются попытки получения гибридных белков, объединяющих части полипептидных цепей белка стержней аденовирусов и белков-лигандов, например пептидных гормонов, в частности с гастрин-релизинг-пептидом. Исследуется возможность подобного изменения белков пентона с тем, чтобы они взаимодействовали с альтернативными тканеспецифическими интегринными. Получены первые результаты, которые указывают на возможность прямой доставки трансгенов с помощью измененных аденовирусов к клеткам эндотелия и некоторым клеткам опухолей (в частности меланомы), которые экспрессируют интегрины $\alpha_v\beta_3$. При этом исключалась возможность взаимодействия аденовирусов с клетками других типов, например эпителия, экспрессирующими интегрины $\alpha_v\beta_5$.

Молекулярные конъюгаты векторной ДНК с лигандами. При таком подходе, используя поликатионы (например полилизин), создают комплексы очищенной векторной ДНК с лигандами. Получены молекулярные конъюгаты векторов с асиалогликопротеинами или IgA в качестве лигандов для доставки конъюгатов к гепатоцитам или клеткам эпителия дыхательных путей соответственно. Основная проблема, с которой приходится сталкиваться на этом пути, заключается в том, что конъюгаты после интернализации попадают в

эндосомы, что ограничивает последующую экспрессию трансгенов. Для преодоления затруднений такого рода в состав конъюгатов пытаются включать компоненты аденовирусов, которые обладают эндосомолитической активностью. Нестабильность комплексов конъюгатов ограничивает их широкое применение в настоящее время.

Использование липосом для направленной доставки трансгенов.

Направленная доставка трансгенов с использованием липосом уже испытана в различных клинических ситуациях. Основным подходом в данном направлении исследований является создание конъюгатов липосом с антителами или лигандами. Например, липосомы, объединенные с антителами к антигенам главного комплекса гистосовместимости мышей, обладают значительно большей эффективностью доставки ДНК к соответствующим клеткам-мишеням, чем липосомы сами по себе. Пептидный лиганд трансферрин был использован для направленной доставки ДНК к клеткам эритроидного ряда костного мозга, экспрессирующим на своей поверхности рецепторы трансферрина. Поскольку одним из основных требований к адресной доставке ДНК в ядра является отсутствие деградации трансгенов, часто используют объединение липосом с вирусными частицами (например вирусом Сендай или его компонентами, в частности F-частицами), которые облегчают доставку и делают ее более безопасной для транспортируемых молекул ДНК. Считают, что конечным результатом этого направления исследований должно быть создание "*суперлипосом*", которые в комплексе с соответствующими антителами будут специфически сливаться с мембранами клеток-мишеней, обеспечивая проникновение рекомбинантных ДНК в клетки с помощью соответствующих вирусных белков с последующим транспортом их в ядра, направляемым сигнальными пептидами ядерного транспорта. Следует упомянуть и об использовании липосом для создания так называемых *виросом* – вирусных частиц, содержащих векторные молекулы нуклеиновых кислот, целиком заключенных в липосомы разнообразной структуры. Такой подход используется для изменения тропизма вирусов с целью направленной доставки векторных молекул при генотерапии.

10.5.2. Управление экспрессией трансгенов в клетках-мишенях

Для того чтобы терапевтическое действие трансгенов реализовывалось в

полной мере, часто бывает необходимо обеспечивать их тканеспецифическую экспрессию в клетках-мишенях на протяжении всей жизни индивидуума. Из-за ограниченной емкости современных векторов, используемых для адресной доставки трансгенов, не представляется возможным применение для регуляции экспрессии энхансеров и сайленсеров, которые в норме чаще всего создают условия для тканеспецифического характера экспрессии генов в их обычном генетическом окружении (подробнее см. раздел 3.2). Из экспериментов с трансгенными животными известно, что для эффективной экспрессии трансгена часто требуется присутствие его интронов. Более того, чтобы предотвратить возмущающее действие регуляторных элементов хромосом клеток-хозяев, необходимо включать в трансген протяженные фланкирующие последовательности. В частности, для достижения полного уровня экспрессии кластера глобиновых трансгенов требовалось фланкировать кластер некодирующей 5'-концевой последовательностью β -глобинового гена и 3'-концевой последовательностью ε -глобинового гена длиной в 20 т.п.о. каждая. Таким образом, для достижения полного уровня экспрессии трансгена необходимо вводить в клетки-мишени фрагменты геномной ДНК длиной в 250 т.п.о. и более, что является трудновыполнимой задачей. Для исключения интерференции эндогенных регуляторных элементов с таковыми трансгенов часто требуется присутствие в генно-инженерных конструкциях специальных регуляторных элементов – инсуляторов (см. раздел 3.2.4), которые должны предотвращать действие гетерологичных энхансеров на трансгены. При этом необходимо иметь в виду, что инсуляторы не оказывают защитного действия на экспрессию трансгенов, не интегрированных в хромосому клетки-реципиента. Таково в самых общих чертах генетическое окружение трансгена, действие которого необходимо учитывать при создании соответствующих генно-инженерных конструкций. Однако имеются и другие, лучше изученные механизмы, которые обеспечивают тканеспецифическую экспрессию трансгенов.

Тканеспецифические промоторы. Обеспечение тканеспецифической экспрессии трансгена может достигаться путем использования в генно-инженерных конструкциях тканеспецифических промоторов, которые содержат регуляторные элементы, обеспечивающие избирательную транскрипцию гена в определенных тканях. В связи с этим повышенный интерес вызывают

промоторы гена регулятора трансмембранной проводимости кистозного фиброза (cystic fibrosis transmembrane conductance regulator – CFTR) и различных генов, специфически экспрессирующихся в мышечных тканях. Установлено, что фрагмент ДНК длиной в 2 т.п.о., фланкирующий ген *CFTR* с 5'-конца, обеспечивает, хотя и на низком уровне, экспрессию трансгенов в клетках дыхательных путей человека и мышей как *in vitro*, так и *in vivo*. Аналогичного эффекта удалось достичь с использованием промотора гена α -актина скелетных мышц и энхансера гена легкой цепи миозина в составе ретровирусных векторов. Эксперименты такого рода приближают то время, когда экспрессия трансгенов в процессе генотерапии будет строго управляемой и тканеспецифической, что является необходимым условием использования генотерапии для лечения заболеваний человека.

Промоторы, активируемые в клетках опухолей. Описано несколько случаев, когда клетки опухолевых тканей обеспечивают более высокий уровень экспрессии определенных генов по сравнению с соответствующими клетками здоровых тканей, что достигается повышением уровня активности их промоторов. Это явилось основанием для разработки одной из стратегий генотерапии рака. Например, многие меланомы характеризуются высоким уровнем синтеза меланина, что связано с активацией промоторов гена тирозиназы и генов, ассоциированных с тирозиназой белков TRP-1 и TRP-2. Эти промоторы, специфически активирующиеся в раковых клетках, были использованы для обеспечения тканеспецифической экспрессии гена-активатора предшественников цитостатиков. Другим примером является раковый эмбриональный антиген (carcinoembryonic antigen – CEA) – специфический маркер опухолевых клеток, локализованный на их поверхности. Экспрессия этого белка регулируется на уровне транскрипции. Регуляторные последовательности его гена также были использованы для избирательной активации предшественников цитостатиков в раковых клетках через активацию соответствующего гена. Аналогичные результаты были получены и с использованием регуляторных элементов онкогенов, усиленно экспрессирующихся в опухолевых тканях. Подробнее о терапевтическом использовании этого подхода см. ниже.

Направленная интеграция трансгенов в определенные генетические локусы. Альтернативным, хотя и более трудоемким, подходом к достижению

избирательной экспрессии трансгенов является их введение сразу под контроль эндогенных регуляторных элементов клеток-мишеней, которые присутствуют в сайте интеграции. Известно, что адено-ассоциированный вирус дикого типа осуществляет сайт-специфическую интеграцию своего генома в хромосому 19 человека и что такая специфичность сохраняется и в ряде генно-инженерных конструкций, полученных с его использованием. Однако векторы на основе ДНК адено-ассоциированных вирусов не могут нести вставку, размер которой превышает 2,5 т.п.о., что ограничивает их практическое использование.

10.5.3. Современные достижения генотерапии онкологических заболеваний

Несмотря на разработку множества новых лекарственных препаратов, направленных на лечение онкологических заболеваний, за последние 30 лет не удалось увеличить число пациентов, проживших более 5 лет после начала заболевания распространенными видами рака, например раком молочной железы или кишечника. Этот факт, выявивший принципиальную неэффективность химиотерапии, стимулировал разработку альтернативных методов лечения онкологических заболеваний, среди которых одним из самых многообещающих подходов является генотерапия.

Многие формы онкологических заболеваний протекают на фоне нарушения регулируемой экспрессии онкогенов или антионкогенов. Причиной такого расстройства на молекулярном уровне чаще всего являются соматические мутации. В связи с этим одним из основных подходов к генотерапии рака является доставка в опухолевые ткани соответствующих неповрежденных генов, что классифицируется как прямая генотерапия, т.е. исправление структуры гена. В случае нарушений экспрессии генов-супрессоров доставляемые трансгены должны кодировать и экспрессировать соответствующие функционально активные белки. В отличие от этого для подавления экспрессии аномально активных онкогенов чаще всего используют антисмысловые РНК или рибозимы.

Другую группу экспериментов, направленных на генотерапию рака, которые исторически были осуществлены раньше прямой генотерапии, можно было бы назвать косвенным воздействием на гены опухолевого роста. При таком подходе в результате направленной доставки трансгенов в опухолевые

ткани последние становятся более чувствительными к конкретным экзогенным воздействиям. Это может достигаться путем доставки гетерологичных генов антигенов главного комплекса гистосовместимости, экспрессия которых делает соответствующие клетки иммуногенными, стимулирующими противоопухолевый иммунитет больного организма. Тот же результат получают при экспрессии в опухолевых клетках генов некоторых цитокинов, фактора некроза опухолей, α - или γ -интерферона.

Векторы, переносящие трансгены HLA-B7 и/или IL-2, уже прошли многочисленные клинические испытания. Ранее на лабораторных животных было показано, что экспрессия на поверхности опухолевых клеток рекомбинантных аллоантигенов главного комплекса гистосовместимости приводит к активации иммунной системы животных против опухолевых клеток и регрессии опухолей. В новейших исследованиях рекомбинантные ДНК были успешно применены для лечения меланом. В использованном подходе рекомбинантный ген, кодирующий аллоантиген главного комплекса гистосовместимости человека HLA-B7, непосредственно вводили в опухоли HLA-B7-негативных пациентов в составе ДНК-липосомных комплексов. Обнадёживающие результаты получены и при клиническом применении этого подхода для лечения пяти больных с меланомами на IV стадии. У одного пациента произошла полная регрессия опухоли после двух инъекций рекомбинантной ДНК, у других наблюдали существенное улучшение состояния. Плазмидную ДНК обнаруживали в опухолях (но не в сыворотке крови) на 3–7-й день после начала лечения, а рекомбинантный белок HLA-B7 присутствовал в биоптатах опухолей у всех пациентов. Это свидетельствовало об эффективной экспрессии рекомбинантных генов в клетках меланомы и непосредственном вкладе белка HLA-B7 в терапевтический эффект генотерапии.

Другим косвенным методом является изменение клеток, таких как лимфоциты, инфильтрующие опухоли (TIL), уничтожающих раковые клетки, путем введения в них трансгенов, кодирующих цитотоксические белковые факторы или гены цитокинов для повышения противоопухолевой активности этих клеток иммунной системы. Для осуществления контроля над количеством остающихся живыми раковых клеток после терапевтических воздействий на них *ex vivo* клетки костного мозга маркируют, как правило, геном устойчивости к неомизину.

Ниже будут рассмотрены еще несколько примеров использования генотерапевтических воздействий для лечения онкологических заболеваний человека. Большинство описанных подходов оказались эффективными на экспериментальных животных и прошли клинические испытания.

Замещение дефектных генов-супрессоров опухолей. Мутации в антионкогене p53 обнаруживают в 50% солидных опухолей у человека. Продукт гена p53 выполняет функции активатора транскрипции других генов, ингибирующих переход нормальных клеток от фазы G₁ к S-фазе клеточного цикла. Внутриклеточный уровень этого белка возрастает в ответ на повреждение геномной ДНК, что сопровождается задержкой клеток в фазе G₁ и репарацией повреждений, а также терминальной дифференцировкой или же, если повреждения ДНК оказываются слишком большими, – *апоптозом* – генетически запрограммированным самоубийством клеток. Инактивация белка p53 ассоциируется с неконтролируемым ростом опухолевых клеток многих типов. Было установлено, что введение функционально активного трансгена p53 в клетки с поврежденным геном белка p53 подавляет рост опухолевых клеток или индуцирует в них апоптоз.

Подавление экспрессии онкогенов. В первой части книги уже упоминалось о том, что гены семейства *ras* представляют собой группу протоонкогенов, наиболее часто активирующихся при опухолевом перерождении клеток. В частности, более чем у 30% клеток аденокарцином легкого и у 80% клеток опухолей поджелудочной железы обнаруживается мутация в онкогене *ras*, что ассоциируется с плохим прогнозом протекания заболевания. Введение в опухолевые клетки трансгенов, экспрессирующих антисмысловую *ras*-РНК, приводит к уменьшению уровня пролиферации клеток и понижению их опухолеродности. Проходят клинические испытания ретровирусных векторов, экспрессирующих антисмысловые РНК гена *ras*. Аналогичный эффект может быть достигнут и в результате применения соответствующих рибозимов.

Введение "генов-самоубийц", или сенсibiliзирующих генов в опухолевые клетки. В большинстве систем, основанных на этом принципе, используют ген тимидинкиназы вируса простого герпеса (HSV-tk). Как уже упоминалось выше, такой ген придает клеткам, его экспрессирующим, чувствительность к *ганцикловиру* (GCV) – лекарству, обычно используемому

для лечения герпеса. HSV-tk обладает способностью фосфорилировать GCV с образованием соответствующего монофосфата, который, в свою очередь, фосфорилируется клеточными киназами до GCV-трифосфата. Последний в процессе репликации включается в строящиеся цепи ДНК, вызывая терминацию синтеза ДНК и гибель клеток. Интересным следствием гибели клеток под действием активированного GCV является распространение летального воздействия на соседние клетки. Таким образом, соседние опухолевые клетки, не содержащие трансгена HSV-tk, тем не менее элиминируются. Это явление получило название *эффекта летального соседства* (neighboring-killing effect), или *эффекта свидетеля* (bystander effect). Проводятся клинические испытания HSV-tk-трансгена в составе ретровирусных векторов в качестве действующего начала при лечении опухолей мозга или их метастазов. В этих экспериментах мышинные клетки, экспрессирующие трансген, непосредственно имплантировали в растущие опухоли. Поскольку перенос трансгенов, опосредованный ретровирусами, ограничивается пролиферирующими клетками, в результате экспрессии трансгена происходит избирательная гибель именно опухолевых клеток. Использование подхода с трансгеном HSV-tk в клиниках дало обнадеживающие результаты, так как при этом само заболевание протекало легче и лечение хорошо переносилось пациентами. Были сконструированы и аденовирусные векторы, экспрессирующие трансген HSV-tk, которые в настоящее время испытываются для лечения раковых заболеваний.

Введение трансгенов, активирующих лекарства, в опухолевые клетки. Бактериальный ген цитозиндезаминазы является одним из примеров гена-сенситизатора. Этот фермент способен превращать нетоксичный для человека 5-фторцитозин в 5-фторурацил – потенциальный цитостатик. Поэтому 5-фторцитозин оказывает летальное действие на клетки, экспрессирующие цитозиндезаминазу, и остается индифферентным по отношению к другим клеткам организма человека. Ген цитозиндезаминазы рассматривается в качестве защитного средства, предохраняющего организм от неконтролируемой пролиферации клеток, измененных в результате генотерапии. Если такая нежелательная пролиферация происходит, клетки уничтожают *in vivo* с помощью 5-фторцитозина, нетоксичного для других клеток.

Терапия предшественниками цитотоксических лекарств,

активируемых продуктами экспрессии трансгенов в клетках-мишенях. При этом подходе пытаются усилить специфичность действия цитостатиков благодаря различиям в уровнях экспрессии трансгенов в клетках разных опухолей. Для реализации такой идеи ген-сенсibilизатор доставляется в нормальные и опухолевые клетки с помощью ретровирусного вектора. При этом ген находится под контролем промотора, обеспечивающего его преимущественную экспрессию в опухолевых клетках, так как именно в них присутствуют необходимые для активации транскрипции белковые факторы. Если ген-сенсibilизатор кодирует цитозиндеаминазу, то 5-фторцитозин будет активироваться только в присутствии этого фермента, синтезирующегося в опухолевых клетках (см. выше). Конструкции с регуляторным доменом гена *CEA* используют для лечения метастазов колоректального рака. Другим примером является рекомбинантный вектор, обладающий промотором онкогена *c-erbB-2*, разработанный для лечения рака молочной и поджелудочной желез. Именно в этих опухолях часто наблюдают сверхэкспрессию такого гена.

Защита стволовых клеток от токсического действия лекарств во время химиотерапии. Ретровирусный вектор, экспрессирующий ген множественной лекарственной устойчивости *MDR1*, был использован при раке яичника с целью придания устойчивости стволовым клеткам организма к цитостатикам, используемым для лечения этого заболевания. Экспрессия данного гена делает клетки нечувствительными к таксолу, даунорубину и многим другим цитостатикам и позволяет сохранять их жизнеспособность и пролиферативный потенциал, необходимые для восстановления жизнеспособности организма после применения интенсивной химиотерапии.

10.5.4. Ближайшие перспективы использования генотерапии

Какие же еще заболевания человека можно рассматривать в качестве ближайшей перспективы для генотерапии? Как упоминалось выше, ретинобластома (онкологическое заболевание, при котором поражаются зародышевые клетки сетчатки глаза) возникает вследствие двух независимых мутационных событий. Первая мутация передается по наследству в виде доминантного признака, а вторая возникает в соматических клетках в процессе развития организма. У 25% пациентов наследуемое нарушение является

делецией, затрагивающей участок хромосомы 13q14, а пациенты с нормальным кариотипом содержат делецию в локусе RB1, обнаруживаемую гибридизацией по Саузерну. Для пациентов с делециями в локусе RB1 характерна предрасположенность и к другим онкологическим заболеваниям. Риск заболевания ретинобластомой детей у родителей с делециями в локусе RB1 может быть существенно уменьшен, если в клетки зародышевой линии родителей ввести недостающий участок генома. Аналогичные результаты могут быть получены при лечении болезни Леша–Нихана методами генотерапии на клетках зародышевой линии. Данный синдром вызывается дефицитом фермента пуринового метаболизма гипоксантин-гуанозинфосфорибозилтрансферазы. У пациентов идентифицированы 17 независимых мутаций в гене, расположенном на длинном плече X-хромосомы и кодирующем этот фермент.

Хорошо изучена также этиология болезни Тэя–Сакса. Для пациентов характерно полное отсутствие активности β -гексозаминидазы, что сопровождается накоплением ганглиозидов или сложных сфинголипидов в лизосомах. Несмотря на разнообразие клинических проявлений болезни обнаружены только две мутации в гене β -гексозаминидазы, одна из которых является вставкой 4 п.о. в экзоне 11, а другая расположена в сайте сплайсинга интрона 12.

Метахроматическая лейкодистрофия характеризуется дефицитом арилсульфатазы А и наследуется по аутосомно-рецессивному типу. Субстратом фермента является гликолипид цереброзидсульфат, который входит в состав миелина. Во время заболевания цереброзидсульфат накапливается в лизосомах, что приводит к прогрессирующей демиелинизации нейронов. Ген арилсульфатазы А локализован на хромосоме 22. Три мутации описаны в аллеле I этого гена: две из них, расположенные в кодирующей части, приводят к замене Tyr-193→Ser, а третья нарушает донорный сайт сплайсинга на границе экзона 2. Таким образом, превентивная генотерапия трех вышеперечисленных наследственных заболеваний потребовала бы точной замены мутантных аллелей на нормальные в клетках зародышевой линии носителей этих мутаций.

В отличие от генотерапии на уровне клеток зародышевой линии

генотерапия с использованием соматических клеток сопряжена с меньшими трудностями и потенциально может быть применена для лечения ряда генетических заболеваний. Одним из существенных преимуществ работы с соматическими клетками является меньшая опасность от внесения в геном соматических клеток мутаций вследствие инсерционного мутагенеза. Это связано с тем, что распространение таких мутаций будет ограничено сравнительно небольшим клоном соматических клеток. Большинство предложений по генотерапии с использованием соматических клеток связано в настоящее время с лечением онкологических заболеваний.

Возможность применения генотерапии, объектом которой будут клетки зародышевой линии человека, ставит ряд этических проблем. Прежде всего, при современном уровне развития генетики неизбежен большой риск внесения наследуемых повреждений в геном человека, и такие повреждения будут носить необратимый характер. Кроме того, с развитием генотерапии будет возникать соблазн использования ее достижений для усиления физических возможностей человека. Генетические и социальные последствия такого рода вмешательства в генетическую природу человека в настоящее время непредсказуемы. Исходя из этого, более безопасной представляется генотерапия с использованием генов соматических клеток человека, и именно здесь успехи генной инженерии возможны в ближайшее время.

10.5.5. Успехи генотерапии в модельных экспериментах

В последнее время получены впечатляющие результаты и по коррекции дефектов на генном уровне с помощью направленного переноса генов в клетки мышей. Одним из таких примеров является успешная генотерапия мышей с синдромом shiverer (дрожание конечностей). У этих мышей имеется мутация в гене, кодирующем основной белок миелина, что фенотипически проявляется в виде тремора и конвульсий. Введение гена дикого типа в клетки зародышевой линии больных мышей приводило к рождению фенотипически здоровых особей. У потомства карликовых мышей, не способных синтезировать гормон роста и развившихся из клеток зародышевой линии, в которые ввели гибридный ген металлотioneина и ген гормона роста крыс, наблюдали восстановление способности к нормальному росту и развитию. Путем введения в клетки зародышевой линии гена дикого типа гонадотропин-рилизинг-гормона удалось

восстановить репродуктивную способность у потомства мышей, родители которых были стерильны из-за отсутствия у них этого гормона и, как следствие, недоразвития гонад. В подобных же опытах введение клонированных β -глобиновых генов человека в клетки мышей с β -талассемией сопровождалось восстановлением нормального фенотипа у клеток красной крови и коррекцией анемии. Для мышей, моделирующих сцепленную с X-хромосомой недостаточность по орнитинтранскарбомулазе, характерны облысение, изменение формы волос и повреждение кожи. Нормальный фенотип восстанавливался у потомства мышей, развившихся из клеток зародышевой линии, в которые был введен ген транскарбомулазы дикого типа. При этом восстанавливалась даже нормальная экскреция оротовой кислоты. Интересные результаты были получены при попытке проведения генотерапии у мышей с синдромом диабета линии NOD (nonobese diabetic), заболевание которых, по-видимому, имеет аутоиммунную этиологию, что обусловлено дефектными генами главного комплекса гистосовместимости. У трансгенных мышей, экспрессирующих нормальные гены, не наблюдали воспалительных реакций в островках Лангерганса панкреатической железы и синдрома диабета.

На мышах моделировали также синдром мукополисахаридоза типа VII, для которого характерно нарушение ступенчатой деградации глюкозаминогликанов в лизосомах. Введение нормального гена β -глюкуронидазы человека таким мышам приводило к коррекции как фенотипа, так и самого биохимического нарушения, приводящего к этому фенотипу.

10.5.6. Проблемы, возникающие в связи с практическим применением генотерапии

Несмотря на впечатляющие успехи генотерапии на модельных животных, в настоящее время имеется ряд принципиальных затруднений, препятствующих широкому использованию метода для лечения заболеваний человека. При проведении генотерапии необходимо выполнять, по крайней мере, следующие условия. Каждый ген, интегрированный в геном человека, должен нормально экспрессироваться и в случае необходимости адекватно отвечать на регуляторные сигналы организма. Интеграция гена в геном не должна сопровождаться инсерционным мутагенезом, нарушающим

функционирование жизненно важных генов, и вообще оказывать неблагоприятное генетическое действие. При этом дефектный ген не должен оказывать отрицательного влияния на функционирование трансгена. Необходимо помнить, что частота интеграции вводимого гена в геном эмбриональных клеток сильно варьирует от клетки к клетке, хотя и несколько возрастает при использовании ДНК в линейной форме, а также при увеличении числа копий генов, вводимых в пронуклеус. Интеграция гена часто сопровождается точковыми мутациями, делециями, дупликациями и перестройками ДНК. Кроме того, интеграция генов в геном, как правило, носит случайный характер, что затрудняет введение гена в требуемый участок генома и может сопровождаться инсерционным мутагенезом. Следствием инсерционного мутагенеза может быть инактивация жизненно важных генов или же превращение протоонкогенов в онкогены.

Ни одна из вышеперечисленных технических проблем, возникающих при практическом использовании генотерапии, не является неразрешимой, однако преодоление этих трудностей требует времени. Реализация международного проекта "Геном человека" даст возможность исследователям использовать на практике от 80 000 до 100 000 генов человека, в том числе и для генотерапии. Все это создает реальные перспективы преодоления многих заболеваний человека, считающихся неизлечимыми.

ГЛАВА 11.ДНК-ДИАГНОСТИКА И ДНК-ТИПИРОВАНИЕ

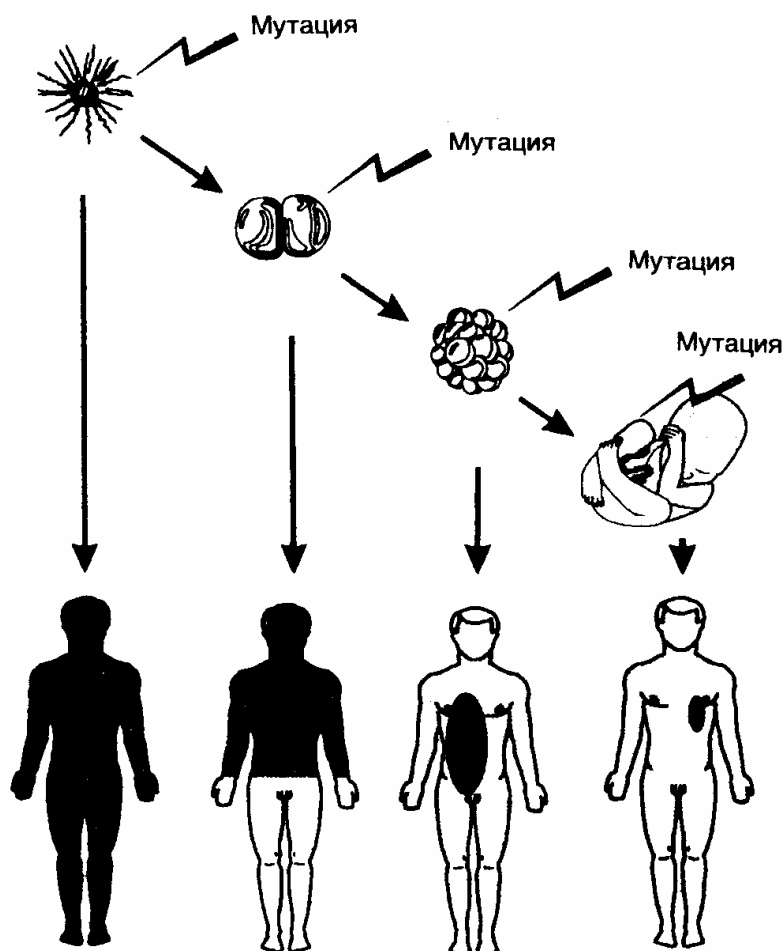


Рис. II.30. Генетические последствия мутаций, происходящих в геноме соматических клеток человека на разных стадиях эмбриогенеза

Черным цветом закрашены места локализации клонов мутантных клеток разных размеров в организме человека

Исследования показывают, что большинство заболеваний человека, а возможно, и все его болезни связаны с мутациями, приводящими к изменению уровней экспрессии конкретных генов. Мутации в определенных генах нарушают функционирование биохимических систем, что в конце концов приводит к развитию соответствующих патологических состояний организма (рис. II.30). В том случае, если мутации происходят в геноме клеток

зародышевой линии человека, все соматические клетки организма-потомка, который развивается из мутантной зиготы, образовавшейся от слияния мутантных гамет в процессе оплодотворения, будут содержать указанную мутацию. Чем позже в онтогенезе человека возникает соматическая мутация, тем меньше размер клона мутантных клеток во взрослом организме. Если же мутация доминантна, т.е. патологический мутантный признак, определяемый мутантным геном, проявляется даже при наличии в соматических клетках копии нормального гена, полученного от другого родителя, то возникает наследственное заболевание.

В случаях, когда мутация, определяющая мутантный фенотип, рецессивна (ее действие проявляется лишь в гомозиготном состоянии, при котором один и тот же мутантный ген получен от каждого из родителей), можно говорить о *предрасположенности* организма-гетерозиготы к соответствующему заболеванию и *носителстве* мутантного гена. Действительно, организм, у которого действие рецессивной мутации маскируется функционированием другого, полноценного аллеля, внешне (фенотипически) выглядит нормальным. Однако у такого организма гораздо больше шансов дать больное потомство в браке с носителем такого же мутантного гена, что является одной из причин запрета на близкородственные браки. С другой стороны, у носителей мутантных генов в гетерозиготном состоянии может произойти соматическая мутация в соответствующем аллельном гене соматических клеток, что также станет причиной развития приобретенного генетического заболевания. В качестве одного из примеров такого заболевания мы уже рассматривали ранее ретинобластому – онкологическое заболевание, поражающее сетчатку глаз, чаще всего, в детском возрасте. Это заболевание развивается в два этапа: вначале организм ребенка получает от одного из родителей мутантный ген *RB1*, являющийся рецессивным онкогеном (антионкогеном), а затем в онтогенезе в одной из соматических клеток в результате мутации происходит инактивация второй его копии, что вызывает малигнизацию клетки и развитие опухоли. Такой же механизм лежит в основе возникновения некоторых форм диабета, а также ряда онкологических и аутоиммунных заболеваний.

Традиционная диагностика наследственных и инфекционных заболеваний строится на детальном изучении симптомов и проведении многочисленных биохимических анализов, включая культивирование

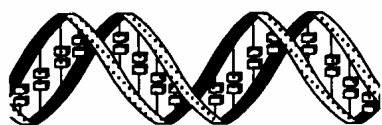
патогенных микроорганизмов на сложных питательных средах. При этом выявление бессимптомного носительства вредных аллелей и микроорганизмов-возбудителей превращается в чрезвычайно трудоемкую задачу. В частности, традиционными цитогенетическими методами можно выявить в геноме человека лишь некоторые крупные хромосомные перестройки: протяженные делеции и транслокации генетического материала, потерю или приобретение целых хромосом. При этом мелкие делеции, транслокации и вставки, а также точковые мутации, которые являются наиболее часто встречающимися изменениями генетического материала, остаются не обнаруженными. То же имеет место и при диагностике бессимптомного носительства возбудителей инфекционных заболеваний бактериальной и вирусной природы, например вируса иммунодефицита человека, которые в латентной стадии инфекции могут присутствовать в организме человека лишь в небольшом числе копий.

Разработанные недавно новые методы молекулярной биологии, и особенно метод ПЦР, значительно облегчили проведение диагностики наследственных и инфекционных заболеваний. При этом мутации выявляются в ДНК клинических образцов, а бессимптомное присутствие возбудителей обнаруживают по наличию геномных ДНК или РНК соответствующих вирусов или бактерий.

В основе всех методов ДНК-диагностики лежат три фундаментальных принципа молекулярной биологии (рис. II.31). Это прежде всего существование комплементарных взаимодействий между двухцепочечными молекулами нуклеиновых кислот (первый принцип), которые после денатурации с последующей ренатурацией (второй принцип) позволяют им безошибочно находить друг друга и восстанавливать первоначальную вторичную структуру. Если же в процессе ренатурации нуклеиновых кислот добавить в пробы короткие олигонуклеотиды (зонды или праймеры), ковалентно соединенные с каким-либо меченым соединением, то они благодаря комплементарным взаимодействиям соединяются с тем участком нуклеиновой кислоты, который содержит последовательность нуклеотидов, строго соответствующую последовательности нуклеотидов зонда или праймера (третий принцип). Следовательно, наличие связавшейся с нуклеиновой кислотой метки с высокой точностью свидетельствует о присутствии в анализируемом образце искомого

последовательностей нуклеотидов.

1. Комплементарные взаимодействия



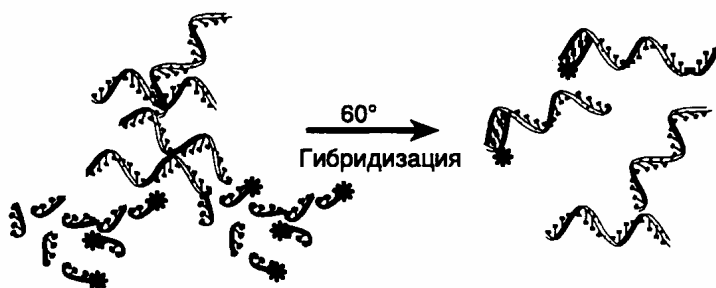
$$A = T$$

$$G \equiv C$$

2. Денатурация–ренатурация ДНК



3. Олигонуклеотидные зонды или праймеры



Длина зонда (праймера) 20–25 нуклеотидов

Рис. II.31. Три молекулярно-генетических принципа, лежащих в основе ДНК-диагностики

Звездочками отмечено наличие радиоактивной, флуоресцентной или другой метки в олигонуклеотидных зондах или праймерах на одном из их концов

Современные методы ДНК-диагностики находят широкое применение в медицине для обнаружения мутаций при наследственных (или приобретенных) заболеваниях и ДНК-типирования организмов (рис. II.32). ДНК-диагностика наследственных и приобретенных заболеваний может заключаться как в обнаружении известных, уже описанных мутаций, так и в поиске новых мутаций, приводящих к тому же патологическому состоянию организма. *ДНК-типированием* называют поиск генотипических особенностей у микро- или макроорганизма, которые позволяют его идентифицировать и(или) отнести к той или иной систематической группе.

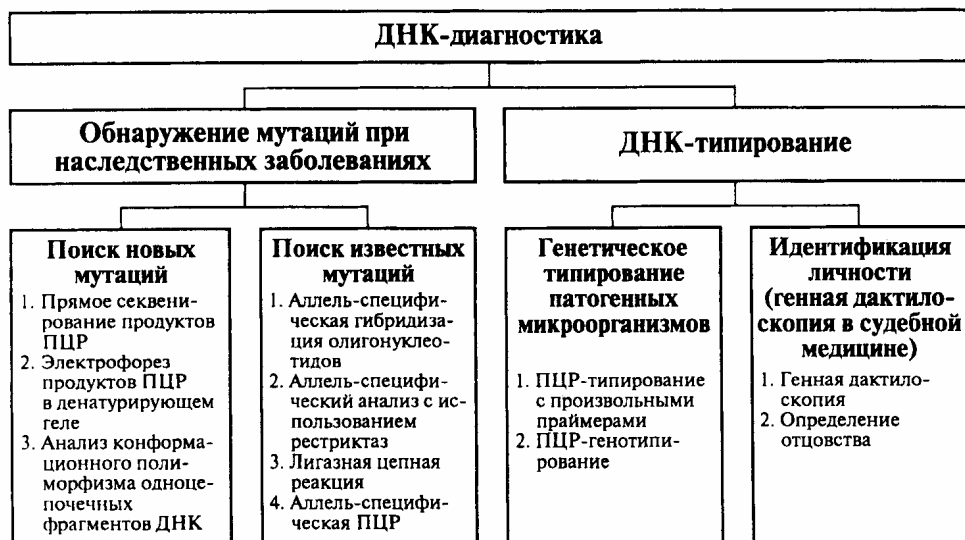


Рис. II.32. Применение методов ДНК-диагностики в медицине

11.1. ДНК-диагностика наследственных и приобретенных заболеваний

Для проведения ДНК-диагностики наследственных и приобретенных заболеваний необходимо подготовить нуклеиновые кислоты, содержащиеся в клинических образцах, и проанализировать их адекватными методами. При этом целью анализа является обнаружение в геноме больного известных или новых мутаций, ассоциированных с развитием мутантного фенотипа в виде симптомов конкретного заболевания.

11.1.1. Получение клинического генетического материала

Для проведения ПЦР используют ДНК клеток различных органов и тканей человека. Основными требованиями, предъявляемыми к такой ДНК, является отсутствие сильной ее деградации и повреждений химическими агентами. Источниками клеток для выделения ДНК служат слюна и другие биологические жидкости пациентов, свежая или высушенная кровь, биоптаты различных тканей, в том числе и корни волос, а также консервированные ткани.

Слюну пациентов (до 5 мл) обычно собирают в пластиковые пробирки. Допускается сбор слюны на фильтровальную бумагу с последующим ее высушиванием с целью облегчения транспортировки. Поскольку источником

ДНК в слюне являются соматические клетки слизистых, а также присутствующие в ротовой полости микроорганизмы, необходимо предпринимать усилия для их механического освобождения с внутренней поверхности щек, например движениями языка или используя специальные щетки. В лаборатории белки слюны расщепляют протеиназой К, а освободившуюся ДНК дополнительно депротеинизируют фенолом и концентрируют осаждением спиртом.

Корни волос содержат несколько сотен соматических клеток, а следовательно, и адекватное число копий ДНК, достаточное для проведения ПЦР. При использовании ДНК из этого источника из-за малого ее содержания необходимо применять особые меры предосторожности против ее загрязнения посторонними ДНК. Преимуществом корней волос перед другими источниками ДНК является почти полное отсутствие внеклеточного матрикса, затрудняющего выделение ДНК из соматических клеток. Это позволяет исключать стадию фенольной обработки образцов для получения ДНК, пригодной для проведения ПЦР. Клетки корней волос лизируют буфером, содержащим неионные детергенты, белки гидролизуют протеиназой К, которую далее инактивируют прогреванием при 100°. При этом одновременно происходит денатурация ДНК, что необходимо для проявления ее матричных свойств при проведении ПЦР.

Срезы для микроскопии и парафинированные ткани используют для проведения семейного генетического анализа, когда другие возможные источники ДНК недоступны из-за смерти пациента или по другим причинам. Для фиксации тканей при проведении патологоанатомических исследований часто используют формальдегид или глutarовый альдегид, которые модифицируют основания ДНК. Поэтому при использовании ДНК из срезов фиксированных тканей возможна амплификация лишь коротких ее сегментов длиной в несколько сотен пар оснований.

Ткани, освобожденные от парафина ксилолом или октаном или же соскобленные с препаратов для микроскопии, суспендируют в лизирующем буфере с протеиназой К и инкубируют в течение 12–24 ч при 56–65° для протеолитического расщепления белков. Дальнейшую депротеинизацию образцов и осаждение ДНК осуществляют так же, как и в предыдущем случае.

11.1.2. Диагностика заболеваний

В процессе диагностики и исследования генетических механизмов наследственных заболеваний человека возникают две тесно связанные друг с другом задачи. На первом этапе исследований в ДНК из клинических образцов пытаются обнаружить уже описанные в литературе мутации, ассоциированные с конкретным наследственным заболеванием. На наличие такой мутации в геномной ДНК пациента указывает диагноз наследственного заболевания, поставленный на основании симптомов болезни. В то же время необнаружение известных мутаций в его геномной ДНК не может считаться доказательством отсутствия у него соответствующего заболевания, поскольку оно может вызываться новыми мутациями в том же самом или других генах, функционально связанных с первым. Поэтому в таких случаях на втором этапе исследования этиологии заболевания следует воспользоваться методами, позволяющими обнаруживать неизвестные мутации.

Методы обнаружения новых неизвестных мутаций. Если раньше точковые мутации, а также микроделеции и вставки обнаруживали преимущественно прямым секвенированием клонированных фрагментов ДНК, то в последнее время преобладающими становятся методы анализа продуктов ПЦР. При этом наиболее часто используются четыре ниже приведенных метода.

Прямое секвенирование продуктов ПЦР является очень эффективным и чувствительным методом выявления мутаций. Принято считать, что с помощью прямого секвенирования одноцепочечных и двухцепочечных продуктов ПЦР стандартными методами можно обнаруживать до 99% мутаций при минимальном количестве ложноположительных и ложноотрицательных результатов. Метод позволяет не только точно локализовать мутацию в исследуемой последовательности нуклеотидов, но и определить ее тип и возможный механизм возникновения с учетом контекста окружающих ее нуклеотидов.

Электрофорез продуктов ПЦР в денатурирующем градиентном геле (denaturing gradient gel electrophoresis – DGGE). Последовательность нуклеотидов в двухцепочечных фрагментах ДНК определяет особенности плавления этих фрагментов в денатурирующих условиях. Отсюда следует, что

мутации (в том числе и одиночные замены нуклеотидов) могут изменять температуру плавления (денатурации) локальных участков ДНК, называемых *доменами*, которые содержат эти мутации. Совокупность доменов плавления определенного участка ДНК образует *профиль плавления*. Домены ДНК в зависимости от последовательности их нуклеотидов в условиях постоянной температуры будут плавиться при определенных концентрациях денатурирующего агента. В результате положение фрагментов в геле по завершении электрофореза будет зависеть от профилей их плавления и, в конечном счете, от нуклеотидной последовательности доменов плавления. Наличие мутации увеличивает или понижает температуру плавления соответствующего домена. Следовательно, фрагмент ДНК, содержащий мутацию, плавится при более низкой (или более высокой) концентрации денатурирующего агента в геле и движется медленнее (или быстрее) при электрофорезе по сравнению с аналогичным фрагментом ДНК дикого типа.

Присутствие ошибочно спаренных нуклеотидов (*mismatches*), как правило, увеличивает разницу в температурах плавления мутантных фрагментов ДНК и фрагментов ДНК дикого типа. С учетом этого факта можно увеличить чувствительность обсуждаемого метода путем сравнения электрофоретических подвижностей фрагментов ДНК дикого типа и гетеродуплексов, образованных фрагментом исследуемой ДНК и соответствующим фрагментом ДНК дикого типа. На практике получение гетеродуплексов продуктов ПЦР не занимает много времени. Для этого достаточно смешать анализируемые продукты в эквимольных отношениях в определенных ионных условиях, денатурировать их повышением температуры и медленно охладить.

С появлением метода ПЦР получение фрагментов ДНК определенной длины, полностью идентичных анализируемому участку ДНК, стало простой задачей. Для повышения разрешающей способности электрофореза в денатурирующем градиентном геле один из праймеров при проведении ПЦР выбирают таким образом, чтобы его GC-состав был максимально возможным. Это предотвращает преждевременную полную денатурацию и расхождение цепей продуктов ПЦР во время электрофореза, так как противоположные цепи денатурирующих фрагментов ДНК будут более продолжительное время связаны друг с другом GC-богатыми последовательностями праймеров.

Полагают, что метод электрофоретического разделения фрагментов ДНК в гелях с градиентом денатурирующих агентов позволяет обнаруживать до 95% мутаций. Однако несмотря на высокую эффективность, этот метод не дает возможности точно локализовать мутации в анализируемых участках ДНК.

Методы, основанные на обнаружении ошибочно спаренных нуклеотидов. Азотистые основания ДНК содержат реакционноспособные кето- и аминогруппы, которые можно легко модифицировать химическими методами. Основания, находящиеся в составе канонической двойной спирали ДНК, в большей степени защищены водородными связями от действия модифицирующих агентов, чем ошибочно спаренные основания в гетеродуплексах. Поэтому последние можно избирательно модифицировать химически.

В настоящее время используют в основном два подхода к обнаружению мутаций с помощью химических модификаций ошибочно спаренных оснований в гетеродуплексах. В одном из них модификации подвергаются ошибочно спаренные остатки цитозина или тимидина с использованием соответственно гидроксилamina или четырехокси осмия. После удаления из цепей анализируемых фрагментов ДНК неспаренных пиримидинов и расщепления депиримидинизированных остатков рибозы по фосфодиэфирным связям пиперидином образовавшиеся фрагменты ДНК исследуют с помощью электрофореза, что позволяет локализовать места расщепления цепей ДНК и соответственно положение мутантных нуклеотидов. Комбинированное использование этих методов дает возможность находить все теоретически возможные сочетания мутантных нуклеотидов.

При втором подходе используют водорастворимый карбодиимид для модификации ошибочно спаренных остатков G и T. ДНК-полимераза прекращает элонгацию строящейся цепи ДНК вблизи модифицированного основания. В результате образуется укороченный гомогенный продукт реакции в виде вновь синтезированного фрагмента ДНК. Поэтому модифицированное основание (а следовательно, и мутация) может быть обнаружено посредством удлинения праймера Taq-полимеразой в присутствии радиоактивно меченных дезоксирибонуклеозидтрифосфатов с последующим электрофоретическим определением размеров меченого продукта реакции. Принято считать, что эффективность выявления мутаций с использованием химических

модификаций ошибочно спаренных нуклеотидов приближается к 95%.

Обнаружение мутаций путем оценки конформационного полиморфизма одноцепочечных фрагментов ДНК (*single-stranded conformation polymorphism – SSCP*). Отдельные цепи денатурированных фрагментов ДНК обладают способностью образовывать вторичную структуру, особенности которой зависят от последовательности нуклеотидов в этих фрагментах. Конформация одноцепочечных фрагментов ДНК, образовавшаяся во время их *рефолдинга* (сворачивания после тепловой денатурации исходного двухцепочечного фрагмента и быстрого охлаждения, во время которого большинство фрагментов остается в одноцепочечной форме), оказывает влияние на электрофоретическую подвижность фрагментов. Наличие в анализируемом фрагменте ДНК, в том числе и в продукте ПЦР, одной или нескольких мутаций приводит к формированию специфической вторичной структуры у цепей анализируемого фрагмента ДНК, образовавшихся в процессе его денатурации с последующим рефолдингом, и эта структура будет отличаться от структуры фрагментов ДНК дикого типа, что легко обнаруживается по изменению электрофоретической подвижности фрагментов. Метод оценки конформационного полиморфизма одноцепочечных фрагментов ДНК прост в постановке и позволяет выявлять до 80% мутаций. Этот метод, однако, не дает возможности точно локализовать мутации.

Еще один недавно разработанный метод поиска мутаций был назван **анализом гетеродуплексов**. Установлено, что наличие в двухцепочечном фрагменте ДНК неспаренных оснований (даже одной пары) изменяет конформацию такого гетеродуплекса, и это проявляется в изменении электрофоретической подвижности фрагмента при проведении гель-электрофореза в определенных условиях. Несмотря на свою простоту, метод также позволяет находить до 80% мутаций.

Все методы, описанные в этом разделе, приобрели особую значимость с развитием метода ПЦР, который достаточно просто позволяет амплифицировать конкретные участки геномной ДНК, получать препаративные количества гомогенной по размерам двухцепочечной ДНК и быстро создавать гетеродуплексы. Несмотря на такое широкое распространение ПЦР, в заключение раздела о методах обнаружения неизвестных мутаций целесообразно упомянуть старый, но эффективный **метод выявления**

мутаций, приводящих к полиморфизму длин рестрикционных фрагментов ДНК (ПДРФ) (*restriction fragments length polymorphism – RFLP*), непосредственно в геномной ДНК. Наличие сайтов рестрикции в геномной ДНК и их взаимное расположение однозначно определяются последовательностью нуклеотидов исследуемой ДНК, поскольку сам сайт рестрикции – не что иное, как строго определенная последовательность нуклеотидов ДНК, узнаваемая и расщепляемая рестриктазами. Следовательно, любая мутация, изменяющая последовательность нуклеотидов сайта рестрикции, уничтожает этот сайт.

Полное расщепление анализируемой геномной ДНК отдельными рестриктазами приводит к образованию определенного набора фрагментов ДНК, число и размеры которых соответствуют расположению сайтов рестрикции. Гибридизация по Саузерну позволяет определять размеры и взаимное расположение рестрикционных фрагментов ДНК после их электрофоретического разделения (подробности см. в разделе 7.5). При наличии мутации в одном из сайтов рестрикции этот сайт остается нерасщепленным после завершения рестрикции, что приводит к слиянию соседних рестрикционных фрагментов ДНК, разделяемых мутантным сайтом, и образованию фрагмента ДНК большего размера. В результате длина рестрикционных фрагментов ДНК, содержащих мутантные сайты, может меняться (становится полиморфной), что четко выявляется при сравнении ДНК из разных источников методом ПДРФ.

Метод ПДРФ широко используется в генетических исследованиях популяций, поскольку наличие в геноме исследуемого организма рестрикционного фрагмента ДНК определенной длины является прекрасным генетическим маркером и одновременно фенотипическим признаком, тесно связанным с генотипом организма. Это позволяет легко следить за распространением такого маркера в популяциях, за передачей его от родителей к потомству при скрещиваниях и использовать для построения генетических карт исследуемых организмов классические генетические методы. ПДРФ-маркеры благодаря их четкой принадлежности определенным генетическим локусам не уступают по информативности распространенным биохимическим маркерам и во многих случаях оказываются удобнее сложных фенотипических признаков (таких, как цвет глаз, окраска волос или шерсти,

форма цветков и листьев), определяемых многими генными локусами.

Методы обнаружения известных мутаций. Все рассмотренные выше методы поиска новых мутаций с успехом используются и для идентификации известных мутаций в исследуемой ДНК. Тем не менее, для выявления уже описанных в литературе мутаций, приводящих к отдельным патологическим состояниям организма (например серповидно-клеточной анемии или нарушениям в системе свертывания крови), бывает удобнее использовать некоторые специфические методы, четыре из которых будут рассмотрены ниже.

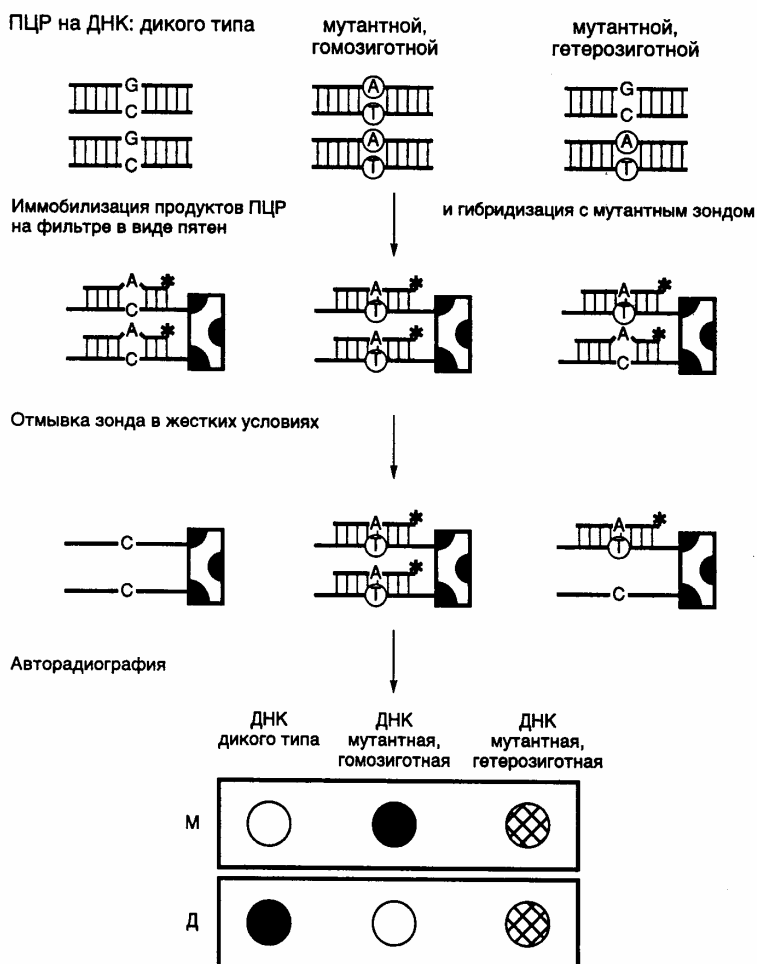


Рис. II.33. Схема обнаружения точковых мутаций гибридизацией с аллель-специфическими зондами

М и Д – зонды, соответствующие мутантному аллелю и аллелю дикого типа. Звездочкой отмечены радиоактивные метки

Аллель-специфическая гибридизация олигонуклеотидов (*allele-specific oligonucleotide hybridization – ASO*). Фрагменты ДНК, амплифицированные с помощью ПЦР с использованием в качестве матрицы ДНК обследуемых индивидуумов и ДНК дикого типа, денатурируют и иммобилизуют на нитроцеллюлозных или нейлоновых мембранах. Мембрану со специфически связанными с ней нуклеиновыми кислотами обычно называют *блотом*. В том случае, если ДНК наносят на мембрану в виде круглых пятен с помощью микропипетки, речь идет о *дот-блотах*. Одни из дот-блотов гибридизуют с олигонуклеотидом (обычно меченным ^{32}P или флуоресцентной меткой), первичная структура которого соответствует исследуемому аллелю дикого типа (*дот-блот-гибридизация*). Другой идентичный дот-блот гибридизуют с меченым олигонуклеотидом, соответствующим мутантному аллелю (рис. II.33). Температуры плавления образовавшихся гетеродуплексов и правильных гибридов различаются. Поскольку в гетеродуплексах имеется ошибочно спаренный нуклеотид, температура их плавления (денатурации) будет ниже, чем у соответствующих правильных гибридов. Подбрав температуру, при которой происходит плавление гетеродуплексов, но не правильных гибридов, можно путем промывки буфером соответствующей температуры избирательно убрать с дот-блотов меченые олигонуклеотиды гетеродуплексов, сохранив олигонуклеотиды правильных гибридов в связанном с ДНК дот-блотов состоянии.

После автордиографии дот-блот-гибридов, промытых буфером соответствующей температуры, получается следующая картина. В случае гибридизации ДНК дот-блотов с олигонуклеотидным зондом, соответствующим последовательности нуклеотидов нормального аллеля, гибрид сохраняется, если в доте присутствует ДНК дикого типа здорового индивидуума. Гибрид разрушается, если в доте присутствует ДНК индивидуума с мутантным геном в гомозиготном состоянии (оба аллеля исследуемого гена, полученные от каждого из родителей, мутантны). В этом случае в результате гибридизации зонда с ДНК обоих аллелей образуются гетеродуплексы, содержащие ошибочно спаренный нуклеотид и обладающие пониженной температурой плавления. В том случае, если обследуемый индивидуум гетерозиготен по исследуемому гену и его ДНК включает в себе как мутантный аллель, так и аллель дикого типа, в результате промывки разрушается гибрид мутантного

аллеля, но не аллеля дикого типа, и половина образованных гибридов сохраняется.

При использовании в дот-блот-гибридации олигонуклеотидного зонда, соответствующего мутантному аллелю исследуемого гена, по тем же самым причинам после промывки в жестких температурных условиях гибриды сохраняются в случае ДНК гомозиготных и гетерозиготных по мутантному аллелю индивидуумов, но будут разрушены в пробах с ДНК, содержащей только нормальные аллели.

Аллель-специфический анализ с использованием рестриктаз применяют в том случае, когда исследуемая мутация изменяет сайт рестрикции. Принцип этого метода полностью совпадает с таковым метода ПДРФ, рассмотренного выше. Рестриктаза может расщеплять двухцепочечную ДНК только в том случае, если ДНК содержит соответствующий сайт рестрикции. Наличие или отсутствие сайта рестрикции в продукте ПЦР указывает на особенности первичной структуры амплифицируемого участка ДНК: наличие или отсутствие мутации, изменяющей последовательность нуклеотидов этого участка. В результате мутаций могут не только элиминироваться известные сайты рестрикции, но и возникать новые, что выявляется по изменению электрофоретической подвижности продуктов ПЦР после их инкубации с рестриктазами.

Аллель-специфическая ПЦР. Известно, что для эффективного осуществления ПЦР 3'-концевой нуклеотид праймера должен быть комплементарен соответствующему нуклеотиду матричной ДНК. В противном случае эффективность удлинения праймера во время ПЦР резко снижается и при определенных сочетаниях ошибочно спаренных нуклеотидов может отсутствовать вообще. Именно эта особенность ПЦР и лежит в основе метода обнаружения мутаций с помощью аллель-специфической ПЦР, иначе иногда называемой *аллель-специфической элонгацией праймера (allele specific primer extension)*.

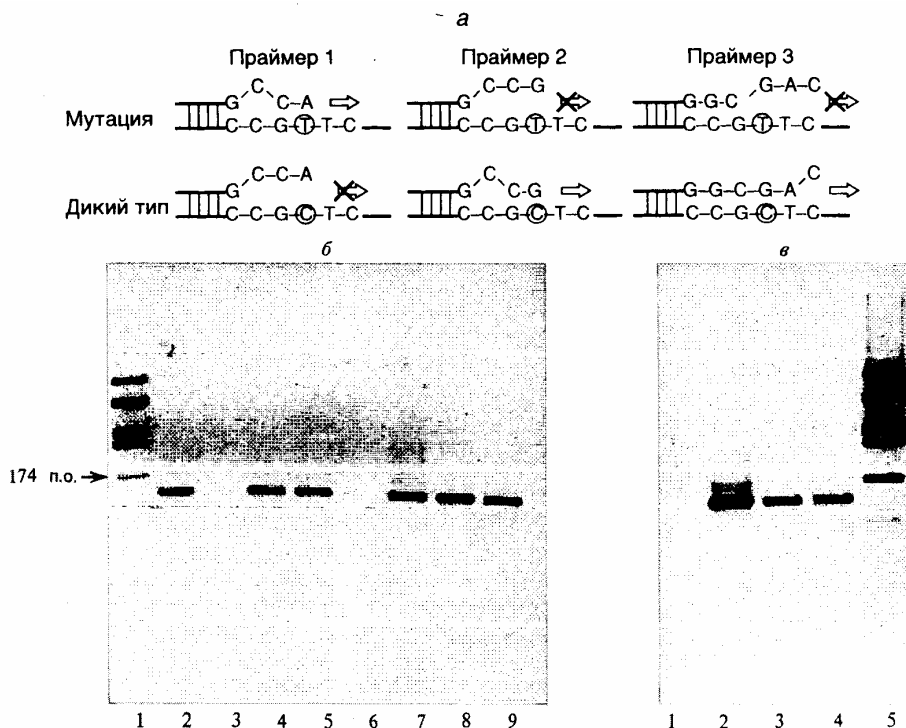


Рис. II.34. Схема одной из современных модификаций аллель-специфической ПЦР (а), использованной для идентификации точковой мутации Leiden в геноме человека (б и в)

Со всеми аллель-специфическими праймерами использован один и тот же встречный праймер. 3'-Концевые нуклеотиды аллель-специфических праймеров комплементарны мутантному нуклеотиду Т (4), нуклеотиду дикого типа С (6), праймер 7 комплементарен нуклеотиду дикого типа своей внутренней частью, а его 3'-концевой нуклеотид не комплементарен матричной ДНК. Для усиления специфичности в праймеры 4 и 6 вблизи 3'-концов введены дополнительные нуклеотиды, некомплементарные матрице. Праймер 4 элонгируется ДНК-полимеразой в присутствии гомозиготной мутантной ДНК, но не ДНК дикого типа (б – дорожки 2 и 3 соответственно), в случае праймеров 6 и 7 наблюдается обратная картина (б и в – дорожки 6,8 и 1,3 соответственно). При наличии мутации в гетерозиготном состоянии все праймеры эффективно элонгируются (б – дорожки 5, 9, в – дорожка 4). Увеличение циклов ПЦР сверх оптимального приводит к образованию продуктов ПЦР в тех пробах, где при меньшем числе циклов они не образовывались (б – дорожки 4,7 и в – дорожка 2), что может служить внутренним контролем. Праймер 7 является универсальным, так как позволяет обнаруживать почти любую гомозиготную точковую мутацию, попадающую в его внутреннюю часть. Использование пар праймеров 4 и 6 или 4 и 7 позволяет легко дифференцировать гомозиготные и гетерозиготные мутации

На рис. II.34 представлены схема высокоэффективного варианта аллель-специфической ПЦР, недавно разработанного нами и использованного для

диагностики мутации *Leiden* при тромбофилиях, а также полученные с помощью данного метода результаты. Эта мутация локализована в экзоне 10 гена фактора V системы свертывания крови человека и часто ассоциирована с синдромом повышенной ее свертываемости. В соответствии с последовательностью нуклеотидов анализируемого участка генома синтезируют два праймера, 3'-концевой нуклеотид одного из которых комплементарен мутантному нуклеотиду матричной ДНК, а у другого – нуклеотиду дикого типа (см. рис. II.34,а,б). Для усиления специфичности действия праймеров вблизи их 3'-концов были введены некомплементарные матрице нуклеотиды. Для исключения ложноотрицательных результатов в пробах, где продукт ПЦР отсутствует, повышали число циклов ПЦР сверх оптимального, после чего, если система работает нормально, продукт ПЦР появляется и в этих пробах, что может служить дополнительным внутренним контролем. Другой тип разработанных нами *универсальных* аллель-специфических праймеров содержит 3'-концевой нуклеотид, всегда некомплементарный матрице, а мутантный нуклеотид матрицы попадает в его внутреннюю часть (см. рис. II.34,а,в). В этом случае продукты ПЦР отсутствуют, если в гибриде во внутреннюю часть праймера попадает любой некомплементарный мутантный нуклеотид матричной ДНК вне зависимости от его точной локализации. Такие праймеры позволяют обнаруживать любые точечные мутации в гомозиготном состоянии и у гаплоидных микроорганизмов.

Для выявления гетерозиготного состояния анализируемого гена иногда используют мутантный и нормальный праймеры разных размеров, которые различаются по длине их 5'-концевых последовательностей, полностью комплементарных анализируемой ДНК. В этом случае в процессе ПЦР в реакционную смесь можно одновременно добавлять оба аллель-специфических праймера: мутантный и нормальный вместе с общим для обоих праймеров – встречным. Образовавшиеся продукты реакции, соответствующие мутантному аллелю и аллелю дикого типа, далее разделяют электрофоретически в гелях, используемых для секвенирования нуклеиновых кислот.

При использовании аллель-специфических праймеров для обнаружения мутаций необходимо иметь в виду, что не все сочетания ошибочно спаренных с матрицей 3'-концевых нуклеотидов одинаково эффективны в блокировании

ПЦР. Наименьшей способностью к элонгации праймеров с ошибочно спаренными 3'-концевыми нуклеотидами обладает так называемый *фрагмент Шоффлера* Taq-полимеразы, который представляет собой молекулу ДНК-полимеразы *T. aquaticus*, укороченную с N-конца на несколько десятков аминокислотных остатков. В то же время в этом методе не применимы термостабильные ДНК-полимеразы, обладающие 3'→5'-корректирующей активностью, например Vent-полимераза.

ПЦР с аллель-специфическими праймерами является простым и эффективным методом обнаружения мутаций в геномной ДНК обследуемых индивидуумов. В отличие от всех рассмотренных выше методов аллель-специфическая ПЦР в такой постановке позволяет находить небольшое количество мутантных ДНК на фоне большого числа молекул ДНК дикого типа. Аналогичная генетическая ситуация может иметь место в том случае, если соматические мутации возникают в процессе онтогенетического развития организмов, и лишь небольшая часть соматических клеток (клон соматических клеток) таких организмов-мозаиков содержит анализируемые мутации. В этом случае, например при онкологических заболеваниях, мутантные ДНК в препаратах суммарной ДНК сильно разбавлены соответствующими последовательностями дикого типа и их трудно обнаружить другими методами. Аллель-специфические праймеры, полностью комплементарные лишь мутантным последовательностям анализируемой ДНК, вовлекаются в амплификацию таких мутантных последовательностей, а последовательности нуклеотидов ДНК дикого типа не амплифицируются. Подобный подход позволяет обнаруживать несколько десятков или сотен молекул мутантной ДНК на фоне десятков тысяч молекул ДНК дикого типа.

Лигазная цепная реакция – ЛЦР (ligase chain reaction – LCR).

Способность ДНК-лигаз восстанавливать фосфодиэфирные связи в одноцепочечных разрывах двухцепочечных молекул ДНК используется при проведении ЛЦР. Для реализации этого подхода синтезируют два олигонуклеотида, комплементарные непрерывной последовательности нуклеотидов ДНК, которые после гибридизации с такой последовательностью примыкают друг к другу и разделены лишь одноцепочечным разрывом. При этом 3'-конец предшествующего нуклеотида должен содержать свободную 3'-ОН-группу, а в 5'-положении 5'-конца олигонуклеотида, следующего за ним,

должна находиться фосфатная группа.

В первом цикле ЛЦР два вышеупомянутых олигонуклеотида отжигаются с исследуемой денатурированной ДНК (как и при ПЦР) в присутствии термостабильной ДНК-лигазы, а также всех остальных необходимых кофакторов и ионов. Реакционную смесь прогревают до температуры, оптимальной для лигирования, после чего между отожденными олигонуклеотидами образуется фосфодиэфирная связь, соединяющая оба олигонуклеотида друг с другом. Затем температуру реакционной смеси повышают до температуры плавления гибрида, образованного анализируемой ДНК и лигированными олигонуклеотидами, которые освобождаются в реакционную смесь в составе единого фрагмента. Во время второго цикла температуру реакционной смеси вновь понижают до температуры отжига олигонуклеотидов, которые, находясь в молярном избытке по отношению к освободившимся продуктам реакции, преимущественно гибридизуются с анализируемой ДНК. Далее связавшиеся олигонуклеотиды снова лигируют, и образовавшийся продукт реакции освобождают из гибрида повышением температуры. При этом содержание продукта в реакционной смеси удваивается. После проведения n циклов ЛЦР в реакционной смеси теоретически может оказаться nx молекул продукта реакции, где x – число пар олигонуклеотидов, вступающих в реакцию в каждом цикле, теоретически равное числу копий анализируемой последовательности нуклеотидов в реакционной смеси.

ЛЦР не происходит в том случае, если 3'-концевой нуклеотид первого олигонуклеотида или 5'-концевой нуклеотид второго некомплементарны соответствующим нуклеотидам анализируемой ДНК. Поэтому при наличии мутации в тех сайтах ДНК, где олигонуклеотиды стыкуются друг с другом, в реакцию со своим партнером вступает только измененный олигонуклеотид, полностью комплементарный анализируемой мутантной ДНК, т.е. продукт реакции образуется лишь при наличии в реакционной смеси "мутантного" олигонуклеотида, но не олигонуклеотида дикого типа. Таким образом, принципы, лежащие в основе метода определения мутаций с помощью ЛЦР, в ряде существенных моментов идентичны принципам аллель-специфической ПЦР.

Для количественной оценки ЛЦР один из олигонуклеотидов обычно метят

биотином, а другой (соответствующий мутантному аллелю) – с помощью радиоактивной или флуоресцентной метки. По завершении ЛЦР олигонуклеотиды, меченные биотином, связывают с мембранными фильтрами с иммобилизованным стрептавидином, а количество лигированных олигонуклеотидов определяют по количеству связавшейся с фильтрами радиоактивной или флуоресцентной метки после отмывки непрореагировавших олигонуклеотидов.

Количественная ПЦР. Результаты большинства из обсуждавшихся выше методов с использованием ПЦР имеют качественный характер. С помощью этих методов легко обнаруживают качественные изменения в генах: точковые мутации, делеции и вставки. Однако при некоторых наследственных заболеваниях соответствующие симптомы возникают в ответ на изменение дозы (числа копий) определенных генов. Типичным примером наследственного заболевания, обусловленного изменением числа копий определенного участка хромосомы 21 в геноме клеток человека, является синдром Дауна. Цитогенетически это заболевание может быть диагностировано по трисомии хромосомы 21 или по наличию крупных транслокаций части этой хромосомы на другие хромосомы на фоне присутствия двух других нормальных хромосом 21. Тем не менее описаны клинические случаи синдрома Дауна, при которых не происходит видимых изменений кариотипа человека. Это обусловлено дупликацией с последующей транслокацией или без нее небольшого участка хромосомы 21. Выявление таких мутаций невозможно традиционными цитогенетическими методами. Однако с помощью количественной ПЦР определение полуторакратного возрастания дозы соответствующего участка хромосомы 21 становится совершенно реальным.

Проведение количественной ПЦР характеризуется, по крайней мере, двумя особенностями. Во-первых, ПЦР проводят в присутствии дезоксирибонуклеозидтрифосфатов или праймеров, меченных с помощью радиоактивной или флуоресцентной метки, с тем чтобы точно определять количество образовавшегося продукта ПЦР. Во-вторых, во время ПЦР необходимо завершать реакцию достаточно рано, пока не образовалось слишком много ПЦР-продуктов. Это связано с тем, что в конечной фазе насыщения, когда продуктов ПЦР становится слишком много, лимитирующими звеньями реакции становятся как субстраты

(дезоксирибонуклеозидтрифосфаты), так и сам фермент, для молекул которого становится слишком много матричной ДНК в виде ПЦР-продуктов. В таких условиях завершение очередного цикла ПЦР уже не сопровождается удвоением количества ПЦР-продуктов и происходит нивелирование небольших количественных различий между разными пробами, которые достаточно четко выявляются на ранних стадиях ПЦР после прохождения нескольких циклов реакции.

При проведении количественной ПЦР в реакционную смесь необходимо вводить внутренние контроли в виде дополнительной пары праймеров, с помощью которых происходит амплификация генетического локуса, для которого точно известна его доза (число копий на молекулу геномной ДНК) в исследуемом геноме. Таким образом, сравнение количества продукта ПЦР, образующегося при амплификации исследуемого локуса, с количеством продукта ПЦР контрольного локуса позволяет определять их количественные соотношения в геноме.

11.2. ДНК-типирование

Результаты, полученные при исследовании структуры и организации геномной ДНК животных, растений и микроорганизмов, наложили глубокий отпечаток на методологию их систематизации. Проблема адекватного отнесения конкретного организма к той или иной группе не является чисто академической. Основанная на точных критериях систематика живых организмов, определяющая эволюционное родство между ними, помимо чисто теоретического представляет и большой практический интерес. В частности, знание источника различных штаммов патогенных микроорганизмов, вызывающих больничные инфекции, позволило бы выработать эффективные меры защиты против их распространения.

Недавно были обнаружены генетические маркеры в виде специфических последовательностей ДНК, которые дают возможность выявлять родственные отношения между особями одного и того же вида путем внутри- и межпопуляционных исследований. Такого рода исследования популяций человека особенно важны с практической точки зрения. В настоящее время молекулярно-генетические методы позволяют осуществлять идентификацию личности в судебно-медицинских исследованиях, а также решать в спорных

случаях проблему определения отцовства.

Генетическим, или ДНК-типированием, называют определение особенностей генотипа организма путем анализа ДНК его генома. Иными словами, в процессе ДНК-типирования определяют особенности первичной структуры ДНК исследуемого организма в конкретных генетических локусах. Абсолютно консервативных генетических локусов, по-видимому, не существует. Мутационные изменения генома непрерывно накапливаются на протяжении филогенетического (исторического, эволюционного) развития вида. Но такого же рода изменения неуклонно происходят в геномах отдельных особей, принадлежащих одной или разным популяциям. Естественно, что проще всего обнаруживаются *межвидовые различия* в последовательностях нуклеотидов исследуемых участков генома, поскольку эти различия, как правило, значительны, и именно они определяют эволюционное расстояние между видами (их *дивергентность*).

Таким образом, у каждого вида организмов имеется большое число внутривидовых различий в первичной структуре ДНК отдельных генетических локусов. Генетические локусы, выполняющие одну и ту же функцию (содержащие один и тот же ген или несколько генов), но различающиеся по первичной структуре ДНК, называют *полиморфными*, а само явление существования в популяции полиморфных локусов получило название *генетического полиморфизма*.

11.2.1. ДНК-типирование микроорганизмов

Наиболее часто в настоящее время используют два способа ДНК-типирования патогенных микроорганизмов, в основе которых лежит метод ПЦР. В первом случае используют один или несколько коротких праймеров произвольной первичной структуры длиной в 6–15 нуклеотидов, которые из-за своих малых размеров обладают низкой специфичностью в отношении конкретных генетических локусов и способны гибридизоваться со многими сайтами геномной ДНК. Во втором случае применяют специфические олигонуклеотидные праймеры длиной в 20–27 нуклеотидов, последовательности которых фланкируют исследуемую последовательность нуклеотидов в бактериальном геноме.

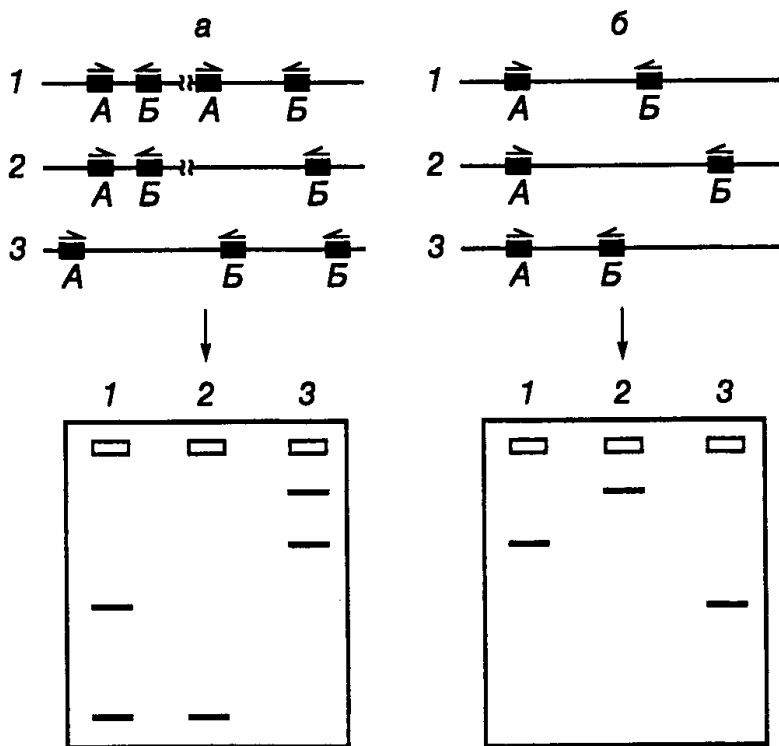


Рис. II.35. Схема ДНК-типирования микроорганизмов с использованием ПЦР и праймеров произвольной структуры

а – штаммоспецифические различия, обусловленные разной локализацией участков ДНК, взаимодействующих с праймерами. В ДНК штаммов 2 и 3 отсутствует участок, который имеется в ДНК штамма 1, что приводит к исчезновению соответствующей полосы продукта ПЦР на электрофореграмме; **б** – результаты амплификации ДНК, содержащей сайты связывания праймеров постоянной локализации. ДНК штаммов 2 и 3 содержат делеции разной длины между сайтами связывания праймеров. Альтернативно, ДНК штаммов 1 и 2 могут содержать вставки, отсутствующие у ДНК штамма 3. Это находит отражение в длинах продуктов ПЦР, разделяемых электрофорезом. А, Б – сайты связывания праймеров на ДНК

Генетическое типирование микроорганизмов с использованием праймеров произвольной первичной структуры. Использование коротких олигонуклеотидных праймеров произвольной структуры основано на том, что в больших геномах для них имеются множественные сайты посадки, а следовательно, и инициации ПЦР. Чем короче такие праймеры, тем большее количество сайтов посадки для них должно существовать. Одним из ограничений дальнейшего участия таких праймеров в ПЦР будет расстояние

между двумя сайтами посадки для противоположно направленных праймеров. Чем длиннее ПЦР-продукт, который должен образовываться в результате функционирования таких праймеров, тем менее надежно работает вся система типирования. На практике длина образующихся с произвольными праймерами продуктов ПЦР находится в пределах 0,2–2,0 т.п.о.

В зависимости от локализации на ДНК мест посадки для пары коротких праймеров могут образовываться два типа продуктов ПЦР. В первом случае различия в длинах продуктов ПЦР обусловлены присутствием на ДНК типизируемых микроорганизмов разного числа сайтов связывания для одного или обоих праймеров (рис. II.35,а). Во втором случае такие различия определяются длинами сегментов ДНК (*ампликонов*), заключенных между парами мест посадки праймеров при неизменном их числе в конкретных генетических локусах (см. рис. II.35,б). На практике могут одновременно реализовываться обе эти возможности. Во время генетического типирования эукариот указанными методами иногда сразу функционируют до 100 ампликонов, тогда как у бактерий это число достигает лишь 20.

Несмотря на то что при обсуждаемом подходе последовательности праймеров выбираются произвольно, получаемая картина амплификации, как правило, является видо- и штаммоспецифичной. При этом количество сайтов связывания на одной и той же ДНК для праймеров одинаковой длины, но с разной первичной структурой может существенно варьировать. На рис. II.35 показаны такие свойства двадцати 10-нуклеотидных праймеров, испытанных на ДНК человека, бобов и *S. aureus*. Видно, что использование праймера AGGGGTCTTG, например, приводит к амплификации 8 ампликонов в ДНК человека, 3 ампликонов в ДНК соевых бобов и не выявляет ни одного ампликона в ДНК *S. aureus*, тогда как при использовании праймера AATCGGGCTG удается обнаруживать до 40 ампликонов в ДНК человека и соевых бобов и до 20 ампликонов в бактериальной ДНК. Таким образом, при использовании обсуждаемого метода в ДНК-типировании самым важным этапом является подбор праймеров или их комбинаций для наиболее эффективного определения принадлежности организма к тому или иному виду, штамму или линии.

Генетическое типирование микроорганизмов с использованием геномных фингерпринтов. При другом подходе к генетическому типированию

с использованием ПЦР исследуют полиморфизм конкретных локусов, для которых хотя бы частично известна первичная структура. Синтезируют специфические олигонуклеотидные праймеры, сайты связывания которых фланкируют исследуемую последовательность ДНК с длиной 1,5–2,5 т.п.о. После проведения ПЦР особенности первичной структуры продуктов ПЦР определяют с помощью рестрикционного анализа. Положение сайтов рестрикции в анализируемой последовательности нуклеотидов может иметь видо- или штаммоспецифический характер и служить точным генетическим маркером того или иного микроорганизма.

С помощью этого метода, который по своей сути является разновидностью рассмотренного выше способа определения ПДРФ, идентифицируют близкородственные, сходные по фенотипу штаммы возбудителей заболеваний, исследуют генетическую структуру популяций микроорганизмов и механизмы их адаптивной изменчивости.

11.2.2. Идентификация личности на основе минисателлитной ДНК: определение отцовства

Определение отцовства представляет собой серьезную социальную, юридическую и медицинскую проблему. Решение этой задачи часто требуется в судах, при разрешении частных споров, для пренатальной (внутриутробной) диагностики, генетических консультаций и при пересадках органов. Например, только в США в 1990 г. было проведено более 120000 тестов на определение отцовства, и это число быстро возрастает. Мировой рынок таких диагностических тест-систем оценивался в 1994 г. более чем в 1 млрд долларов и сейчас является крупнейшим рынком среди молекулярно-генетических диагностик.

С использованием диагностических тест-систем на основе минисателлитных ДНК определение отцовства получило прочную научную основу. С этой целью в настоящее время используют два подхода. При одном из них применяют олигонуклеотидные зонды, специфичные в отношении многих минисателлитных локусов, при другом – наборы зондов (или праймеров), специфичных в отношении отдельных полиморфных локусов VNTR (подробнее о VNTR см. раздел 1.3.1).

Теоретические аспекты возможности определения отцовства.

Возможность определения отцовства, так же как и отнесение биологических образцов, содержащих ДНК, к тому или иному человеку, основано на наличии в геноме человека четких генетических маркеров в виде определенных последовательностей ДНК, набор которых уникален для конкретного индивидуума. Теоретически такими маркерами могли бы быть любые последовательности нуклеотидов ДНК, для которых характерна большая изменчивость в популяциях человека. Как было уже отмечено выше, tandemно повторяющиеся последовательности минисателлитов (VNTR) являются одними из наиболее полиморфных последовательностей нуклеотидов в геноме человека. Поэтому неудивительно, что именно они были использованы для идентификации личности. Наличие сочетания генетических маркеров среди VNTR, общих для мужчины, женщины и спорного ребенка, в ряде случаев может однозначно указывать на родственные связи между ними.

Отсутствие общих генетических маркеров у обследуемых ребенка и мужчины однозначно исключает последнего как отца ребенка. Однако обнаружение у них общих маркеров еще не может быть доказательством того, что подозреваемый мужчина является отцом. Доказательства отцовства основываются на простых статистических расчетах, в которых учитываются частоты встречаемости в популяции общих аллелей исследуемых VNTR-локусов. Рассчитывается отношение (X/Y) вероятности (X) получения наблюдаемого набора маркеров возможного настоящего отца к вероятности (Y) обнаружения этого набора маркеров у любого, выбранного наугад человека, принадлежащего этой популяции. Такое отношение вероятностей получило название *индекса отцовства* (paternity index – PI).

При расчетах PI возникает еще одна проблема, связанная с определением вероятности (X) того, что обследуемый мужчина является отцом. Значение такой вероятности необходимо иметь до проведения каких-либо диагностических тестов. Обычно используемое значение 0,5 нельзя считать достаточно обоснованным во всех конкретных случаях. К счастью, при очень больших значениях PI ($>10^4$), которые обычно получаются при таких исследованиях ДНК, выбор этой исходной вероятности оказывается практически несущественным. Как правило, это получается вследствие низких значений вероятности Y .

Определение отцовства с использованием олигонуклеотидных

зондов, специфичных в отношении нескольких VNTR-локусов. В 1985 г. И.О. Джеффрисом и соавторами впервые было показано, что олигонуклеотидные зонды, комплементарные последовательностям миоглобинового гена человека, одновременно обладают способностью гибридизоваться по Саузерну с множественными локусами минисателлитной ДНК. Профили гибридизации оказались специфичными для отдельных индивидуумов. Совокупность электрофоретически разделяющихся рестрикционных фрагментов анализируемой ДНК, выявляемых после проведения гибридизации с мечеными зондами, которые специфичны в отношении полиморфных минисателлитных локусов, получила название *ДНК-фингерпринтов, или генетических отпечатков пальцев*. С помощью таких и других аналогичных олигонуклеотидных зондов удается выявлять на одной электрофореграмме до 15–20 различных фрагментов ДНК одного индивидуума, молекулярная масса которых превышает 3,5 т.п.о., а также много более мелких фрагментов, которые не учитываются при определении отцовства этим методом.

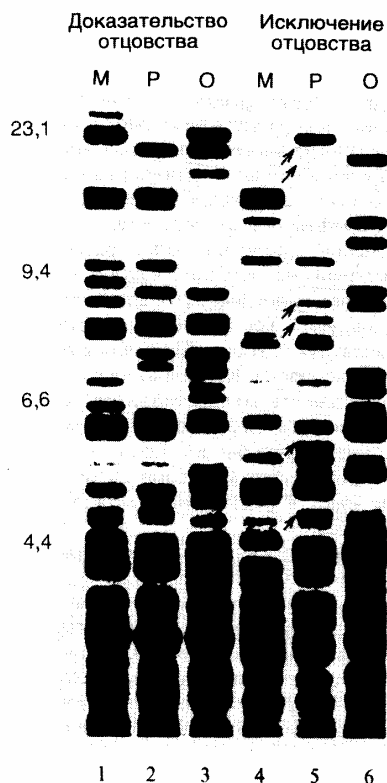


Рис. II.36. Примеры исключения и доказательства отцовства с помощью ДНК-типирования

Результаты гибридизации по Саузерну с зондом F10 показывают полную идентичность фрагментов ДНК у ребенка и отца (дорожки 2 и 3), что рассматривается как доказательство отцовства, либо выявляют, по крайней мере, 6 дополнительных фрагментов ДНК у ребенка, обозначенных стрелками, которые отсутствуют у отца (дорожки 5 и 6 – исключение отцовства)

На рис. II.36 показаны результаты одного из таких опытов. Интерпретация ДНК-фингерпринтов, полученных при анализе множественных локусов, основана на трех постулатах. Прежде всего, предполагается, что фрагменты ДНК, видимые на фингерпринтах, являются аллельными продуктами отдельных генетических локусов человека и передаются потомству независимо друг от друга. Во-вторых, считается, что для каждого генетического локуса частоты встречаемости в популяции отдельных аллелей следуют нормальному распределению Пуассона. И, наконец, принимается, что фрагменты ДНК, электрофоретическая подвижность которых совпадает, представляют один и тот же аллель конкретного локуса. Накопленный опыт работы с ДНК-фингерпринтами показывает, что первое допущение соблюдается достаточно хорошо: *аллелизм* (парность гомологичных генов, определяющих разные фенотипические признаки у диплоидных организмов) и генетическое сцепление между исследуемыми локусами наблюдаются редко. Невыполнение второго предположения не сказывается серьезно на результатах тестирования, поскольку выводы делаются без учета частоты встречаемости отдельного аллеля на основе совпадения структуры (фрагментов ДНК) многих локусов. Третий постулат является более спорным, однако его применение придает значениям индекса отцовства стабильность.

С этими исходными условиями статистическая оценка ДНК-фингерпринтов множественных локусов основывается только на одном параметре: средней доле фрагментов ДНК (x), которые совпадают у людей без родственных связей. Такой параметр в большей степени зависит от техники лабораторных исследований, чем от свойств обследуемой популяции. Это прежде всего способность используемой системы к электрофоретическому разделению индивидуальных фрагментов ДНК (т.е. разрешающая способность используемого метода), принципы выбора конкретных фрагментов ДНК для

анализа, а также критерии принятия решения об идентичности сравниваемых фрагментов ДНК. Следовательно, параметр x может варьировать при сравнении результатов, получаемых в разных лабораториях, но эти различия будут постоянно сохраняться для различных популяций и субпопуляций. Действительно, при использовании, например зондов 33.6 и 33.15, оказалось, что x один и тот же у неродственных индивидуумов, в парах муж–жена и в различных этнических группах.

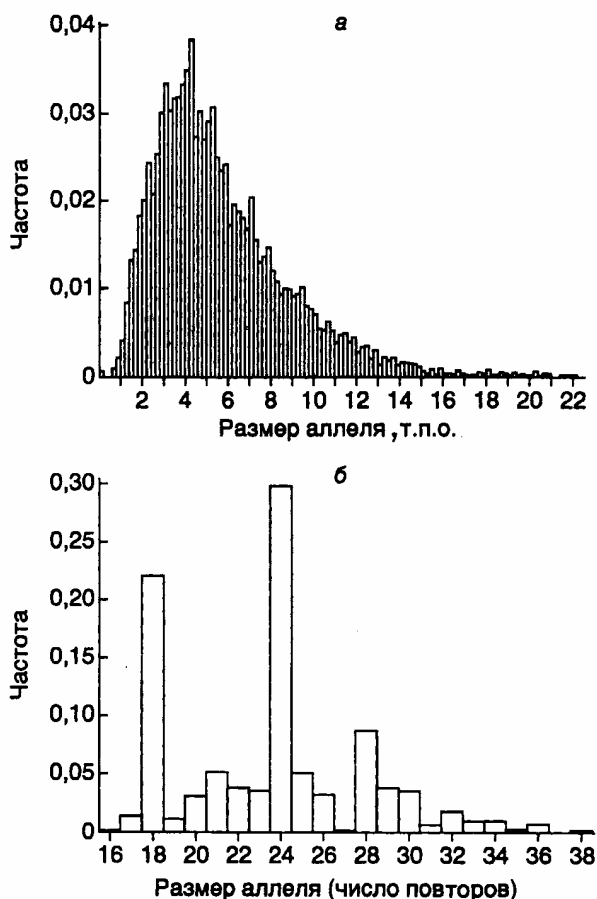


Рис. 11.37. Сравнение информативности двух минисателлитных зондов при идентификации личности

Частоты встречаемости минисателлитных локусов D1S7 (а) и D1S80 (б) определенного размера в популяции оценивали гибридизацией по Саузерну после расщепления ДНК рестриктазой *HaeIII* (а) или ПЦР (б). Для локуса D1S7 характерно квази-непрерывное унимодальное распределение, тогда как локус D1S80 характеризуется меньшей гетерозиготностью (84%) с небольшим числом дискретных аллелей, для которых характерно мультимодальное распределение

Определение отцовства с использованием ДНК-зондов, специфичных в отношении только одного локуса. ДНК-фингерпринты, получаемые при одновременном исследовании многих локусов, отражают скорее фенотип индивидуума, чем его генотип. Действительно, получаемая в итоге картина включает в себе множество полос ДНК, в том числе и неразделившиеся, а также слабо разделившиеся фракции. Такая электрофореграмма напоминает сложный фенотипический признак, например форму лица человека, которая является результатом экспрессии громадного числа генов. В отличие от этого с помощью ДНК-зондов, представляющих собой клонированные последовательности минисателлитов и взаимодействующих только с одним локусом, можно получать истинную информацию о генотипе изучаемого организма. Таким образом, имея дело с отдельными полиморфными локусами человека, исследователи получают в свои руки систему *кодминантных аллелей* (т.е. аллелей, *совместно* участвующих в формировании фенотипических признаков), наследуемых по законам Менделя. Именно понимание наследования таких минисателлитных локусов и привело к широкому распространению однолокусного метода определения отцовства.

В настоящее время получены сотни клонов минисателлитной ДНК и на их основе разработаны комбинации зондов, пригодные для определения отцовства. При выборе зондов для таких минисателлитных локусов обычно руководствуются следующими критериями: зонды должны обладать строгой локус-специфичностью, а тестируемые локусы быть несцепленными (передаваться потомству независимо друг от друга) и обладать достаточной, но не чрезмерной генетической стабильностью. В частности, среди наиболее широко используемых зондов MS1 (D1S7) соответствует генетическому локусу, гетерозиготному в 99% случаев, однако он мутирует с очень высокой частотой (0,05 на гамету) и поэтому, несмотря на высокую информативность, не используется при определении отцовства (рис. II.37,а). В то же время для локуса D1S80 со значительно меньшей вариабельностью (84% гетерозигот) характерно образование кластеров в частотах распределения фрагмента ДНК по длине (см. рис. II.37,б). Поэтому небольшие ошибки в определении длины аллелей могут привести к значительным ошибкам в оценке частоты их встречаемости. Такие локусы достаточно легко изменяются в результате

генетического дрейфа и инбридинга.

В настоящее время разработана теория, указывающая на то, сколько локусов должно быть типировано для получения правильного ответа об отцовстве при известных значениях гетерозиготности этих локусов в популяции. Например, если используются локусы, гетерозиготность которых составляет 90%, для установления отцовства необходимо проанализировать шесть таких локусов. В США в настоящее время для этих целей обычно используется набор из трех–пяти однолокусных зондов. Однолокусные зонды обладают низкой разрешающей способностью относительно братьев и сестер (*сисбсов*). В частности, с помощью одного такого зонда можно лишь с вероятностью 75% обнаружить генетические различия между сибсами, и эта вероятность увеличивается до 99,6% при использовании четырех зондов. Данный факт приобретает особую важность, когда при определении отцовства необходимо сделать выбор между братьями.

Особенности определения отцовства по отдельным локусам с использованием ПЦР. В качестве альтернативы однолокусным зондам в последнее время для определения отцовства часто используют ПЦР. Отдельные мини- и микросателлитные локусы могут быть амплифицированы с помощью праймеров, комплементарных уникальным последовательностям ДНК, фланкирующим эти повторяющиеся последовательности. При идентификации личности метод ПЦР, который по своей сути является одной из разновидностей однолокусной методики, поскольку имеет дело с отдельными локусами, обладает, по крайней мере, двумя преимуществами перед однолокусными зондами. Во-первых, популяционный полиморфизм длин ДНК аллелей, исследуемых с использованием этого метода, носит более дискретный характер, чем у аллелей, изучаемых с помощью однолокусных зондов (см. рис. II.37,а,б). Это обстоятельство облегчает последующее вычисление индекса отцовства. Во-вторых, метод ПЦР обладает гораздо большей чувствительностью и может быть использован при анализе образцов, содержащих <1 нг геномной ДНК и полученных из разных источников (см. раздел 11.1.1).

К недостаткам ПЦР в применении к определению отцовства следует отнести низкую информативность полиморфных микросателлитов и коротких минисателлитов. Это связано с тем, что они обладают <90% гетерозиготности,

небольшим числом аллелей и имеют тенденцию к образованию кластеров по размерам (см. рис. II.37,б). Кроме того, на распределение таких последовательностей в геноме оказывают влияние инбридинг и принадлежность индивидуумов к определенным этническим группам. Количество локусов минисателлитов, которое необходимо исследовать методом ПЦР для определения отцовства, приближается к 11, а микросателлитов – к 18.

Для типирования с помощью ПЦР наиболее часто используются три минисателлитных локуса человека: в гене аполипопротеина В (APOB), D17S5 (известный также как локус D17S30) и D1S80. Все три локуса легко амплифицируются (максимальный размер их аллельных вариантов не превышает 1 т.п.о.) и легко обнаруживаются с помощью электрофореза. Однако для них характерны низкий уровень гетерозиготности и малая изменчивость (что выражается в небольшом числе известных аллелей). Мутации в этих микросателлитах возникают очень редко.

Одним из путей повышения информативности полиморфизмов микросателлитов при ДНК-типировании является одновременная амплификация двух тесно сцепленных микросателлитных локусов, сочетания которых формируют множество *гаплотипов*. Например, одновременная амплификация двух GATA-повторов, локализованных в интроне 40 гена фактора фон Виллебранда, которые разделены последовательностью длиной в 212 п.о., обнаружила суммарный уровень гетерозиготности объединенного локуса ~93%. При этом уровни гетерозиготности индивидуальных локусов составляли лишь 72 и 78% соответственно.

В заключение необходимо еще раз отметить, что по своей логике современные методы определения отцовства, основанные на ДНК-типировании, несколько противоречивы. Если отрицательное заключение об отцовстве, основанное на несовпадении аллелей анализируемых мини- или микросателлитных локусов, абсолютно и не подлежит сомнению, то положительный вывод может быть сделан лишь с некоторой долей вероятности, которая основана на частоте встречаемости конкретных аллелей анализируемых локусов в популяции. С другой стороны, в результате методических ошибок легко могут быть сделаны ложноотрицательные выводы, однако положительное заключение об отцовстве в результате лабораторной

методической ошибки практически исключено.

Высокая информативность многолокусных ДНК-фингерпринтов подтверждена большим числом генетических и популяционных исследований. Эмпирические данные, полученные при обследовании тысяч семей, показали, что с помощью многолокусных зондов можно разрешать все спорные случаи отцовства. Использование однолокусных зондов затруднено невозможностью создания полной классификации соответствующих аллелей в популяции из-за кажущегося непрерывного распределения их по размерам. Однако и в этом случае на практике проблема полностью решается с помощью набора из пяти–шести однолокусных зондов. Применение ПЦР ограничивается большой эволюционной консервативностью амплифицируемых мини- и микросателлитных локусов и, как следствие, малым суммарным числом аллелей. Однако ПЦР бывает очень полезна на первых этапах исследования из-за методической простоты постановки опытов, особенно в условиях малой доступности исходного биологического материала. Все три группы методов хорошо дополняют друг друга и в разных сочетаниях в спорных случаях позволяют однозначно идентифицировать личность человека.

11.3. Микроматрицы и микрочипы ДНК

Одним из интенсивно развивающихся направлений биотехнологии нуклеиновых кислот в последнее время становится использование микроматриц ДНК для анализа нуклеотидных последовательностей. В этой группе методов на небольшой по размеру поверхности стекла или другого твердого носителя иммобилизуют в виде правильно расположенных микропятен небольшие фрагменты ДНК с известной последовательностью нуклеотидов (чаще всего синтетические олигонуклеотиды или кДНК), которые далее используют для гибридизации с анализируемыми образцами нуклеиновых кислот. При совпадении первичной структуры ДНК микропятна и анализируемого образца на поверхности стекла образуются правильные ДНК–ДНК-гибриды, которые обнаруживаются по появлению в данных участках микроматрицы сигналов, например в виде микроскопической флуоресцирующей точки или по тушению флуоресценции исходно меченых последовательностей. Такие кусочки твердого носителя с нанесенными на них микроматрицами ДНК получили название *микрочипов ДНК*.

Для нанесения нуклеиновых кислот на поверхность подложки в основном используют три подхода: короткие олигонуклеотиды синтезируют прямо на ее поверхности, а также прикрепляют к ней предварительно полученные фрагменты ДНК ковалентными или нековалентными связями.

Во время синтеза олигонуклеотидов непосредственно на поверхности стекла применяют те же реагенты и проходят те же стадии, что и при обычном твердофазном синтезе в современных автоматических синтезаторах. В наиболее распространенном варианте при создании микроматрицы используют фотолитографическую маску, которая избирательно закрывает от света и оставляет открытыми участки микроматрицы с синтезируемыми олигонуклеотидами, содержащими чувствительные к облучению светом защитные химические группы. На каждом этапе синтеза маску, которая на схеме обозначена серыми прямоугольниками, помещают над большей частью микроматрицы, а остающиеся открытыми химические группы активируют светом. Затем происходит соединение 5'-гидроксильных групп фосфорамидитных производных добавляемого нуклеотида с активированным сегментом микроматрицы.

Защищая поверхность микроматрицы от света дискретными участками в различных простых комбинациях, удастся синтезировать, как это показано на схеме (рис. II.38), за $4 \cdot 3 = 12$ отдельных этапов все теоретически возможные последовательности тринуклеотидов в количестве $3^4 = 81$. В общем случае для синтеза всех возможных последовательностей длиной в N нуклеотидов требуется $4N$ синтетических стадии. Длина цепи олигонуклеотида, который может быть синтезирован на микроматрице, лимитируется выходом готового продукта на каждой стадии, который составляет $\sim 95\%$. Так, суммарный выход 25-звенного олигонуклеотида будет составлять всего $\sim (0,95)^{25} = 28\%$. В этой связи рассмотренный метод обычно используют для синтеза олигонуклеотидов, длина которых не превышает 20–25 оснований.

При альтернативном подходе микроматрицы олигонуклеотидов синтезируют на подложке с использованием технологии струйных принтеров. В этом случае головка принтера движется вдоль подложки и, в соответствии с заложенной программой, наносит на необходимые участки небольшие количества раствора с фосфорамидитными производными нуклеотидов из индивидуальных резервуаров. Этапы снятия защитных групп и промывания

производятся так же, как и при обычном твердофазном синтезе олигонуклеотидов. Эффективность каждого этапа синтеза в этом случае может превышать 99%, что позволяет синтезировать олигонуклеотиды длиной до 40 оснований с суммарным выходом ~67% (40 этапов с выходом 99% на этап).

Еще одной разновидностью методов создания микроматриц является синтез индивидуальных олигонуклеотидов, их очистка и нанесение с помощью микроробота на поверхность подложки с адгезивным покрытием. Однако создание этим способом микроматриц, содержащих тысячи индивидуальных элементов, – очень трудоемкий процесс.

При конструировании микроматриц, элементы которых содержат индивидуальные кДНК, для нанесения на подложку микропятен чаще всего используют микророботы. В этом случае применяют растворы рекомбинантных кДНК длиной 0,5–1,0 т.п.о., очищенных из бактериальных клеток, которые перед применением как правило, амплифицируют с помощью ПЦР. Поверхность стеклянной подложки покрывают тонким слоем полилизина или обрабатывают аминосиланом с целью создания на ней положительного заряда, что обеспечивает возможность электростатического взаимодействия подложки с отрицательно заряженными молекулами кДНК. Недостатком этого способа является неспецифичность электростатических взаимодействий, в которые вовлечены многие участки кДНК, что уменьшает эффективность взаимодействия кДНК с последовательностями анализируемых образцов. Для преодоления этих затруднений с помощью асимметричной ПЦР синтезируют производные кДНК, которые далее ковалентно соединяют с силированным стеклом с помощью боргидрида.

Разрабатываются и альтернативные конфигурации микроматриц, в которых на поверхности стекла фиксируют сами фрагменты ДНК, анализируемой с помощью олигонуклеотидных зондов. Такой подход был использован, в частности для поиска мутаций в гене белка p53. Дальнейшим расширением этого подхода является нанесение на поверхность стекла кусочков тканей с последующим анализом содержащихся в них ДНК или РНК гибридизацией *in situ*.

11.3.2. Ограничения в использовании микроматриц ДНК

Помимо самой достаточно сложной технологии производства микроматриц, к числу факторов, ограничивающих их широкое применение, относятся кинетические параметры гибридизации, а также точность и чувствительность обнаружения в гибридах ошибочно спаренных нуклеотидов.

Кинетика гибридизации. Гибридизация в гетерогенных смесях нуклеиновых кислот происходит медленно. При использовании олигонуклеотидов в качестве зондов для ее завершения обычно требуется несколько часов, тогда как в случае зондов кДНК процесс продолжается в течение 6–24 ч при высокой температуре. Поиск зондом своей иммобилизованной мишени происходит двумя путями. Во-первых, может произойти прямая гибридизация непосредственно из раствора, эффективность которой в первом приближении не зависит от размера зонда. Во-вторых, зонд вначале может неспецифически связаться с поверхностью подложки, диссоциировать из комплекса и, диффундируя вдоль поверхности, достичь мишени. Скорость гибридизации по второму механизму резко возрастает, когда длина зонда становится меньше 1000 нуклеотидов. С учетом этого время, необходимое для гибридизации, может быть уменьшено путем укорачивания молекулы зонда, повышения его концентрации в гибридизационных смесях и сокращения расстояния, которое необходимо преодолеть зонду до мишени путем диффузии.

Поскольку концентрация зондов в гибридизационных смесях ограничена их химическими свойствами, основным фактором, с помощью которого можно уменьшать время гибридизации и объем наносимого образца с зондом, становится плотность расположения пятен ДНК-мишеней (элементов) в микроматрицах. Размеры индивидуальных элементов микроматриц, получаемых современными фотолитографическими методами, лежат в пределах 5–10 мкм при общей их максимальной плотности на подложке $\sim 10^6/\text{см}^2$. Плотность элементов в обычно используемых микрочипах, полученных с помощью фотолитографии, составляет 10^4 – $10^5/\text{см}^2$. При использовании микророботов для нанесения кДНК на поверхность подложки плотность расположения элементов достигает лишь $\sim 10^3/\text{см}^2$. И в том, и в другом случае зондам для гибридизации необходимо диффундировать к

мишени, находящейся в индивидуальном элементе микроматрицы, общие размеры которой как правило составляют 1 x 1 см. Поскольку время диффузии зонда пропорционально квадрату расстояния до мишени, общее время гибридизации можно понизить путем увеличения плотности расположения элементов на микрочипе. Кроме того, одним из подходов к уменьшению времени гибридизации является придание молекулам зондов направленности перемещения к мишеням в электрическом поле, когда микроматрицу располагают на электроде.

Чувствительность метода. Чувствительность метода специфического обнаружения последовательностей ДНК с использованием технологии микроматриц определяется отношением количества зонда, связавшегося со специфической мишенью, к количеству неспецифически связавшегося зонда. Чем выше это отношение, тем больше чувствительность метода. Неспецифическое связывание зонда может происходить как с самой подложкой, так и с гомологичными зонду последовательностями нуклеотидов ДНК-мишени с образованием неправильно спаренных оснований. В последнем случае могут быть получены ложноположительные результаты. Поскольку площадь пятна специфической мишени значительно меньше доступной зонду поверхности микрочипа, при малых временах гибридизации количество неспецифически сорбированного зонда будет значительно превышать количество зонда, находящегося в правильных гибридах. Эту ситуацию улучшают путем увеличения времени гибридизации до достижения процессом стационарной фазы. В этой связи интерпретация картин гибридизации требует постановки адекватных контролей для оценки эффективности неспецифического связывания зонда поверхностью микрочипа. Кроме того, для количественной оценки гибридизации необходимо использовать количества зонда, не насыщающие мишени микрочипа во время отжига.

Эффективность детекции флуоресценции как таковой является одним из самых сильных факторов, сдерживающих развитие технологии микрочипов ДНК. По мере уменьшения размеров элементов микроматриц все меньшее количество молекул ДНК-мишени становится доступным зонду. Современные методы позволяют обнаруживать ДНК-мишени в аттомолярных концентрациях (10^{-18} М) при плотности в одну молекулу ДНК на 1 мкм^2 . Для этой цели чаще всего используются системы, основанные на конфокальной сканирующей

лазерной микроскопии. При определении альтернативных состояний гомологичных последовательностей ДНК-мишеней, в частности, аллельного (нормального или мутантного) состояния конкретного гена, применяют двухцветные системы. В таких системах один из зондов, например специфичный в отношении аллеля дикого типа, метят флуоресцентным красителем одного цвета (красным), а другой (мутантный) – флуорофором другого цвета (зеленым). Зонды одновременно гибридизуют с анализируемым микрочипом, который далее сканируется при обеих длинах волн, что позволяет обнаруживать тот или иной зонд, сохранившийся в гибриде, и на этом основании делать вывод об аллельном состоянии гена.

11.3.3. Использование микроматриц ДНК в фундаментальных и прикладных исследованиях

Определение первичной структуры и картирование ДНК являются основными направлениями использования микроматриц олигонуклеотидов в настоящее время. Прямое секвенирование генов с помощью олигонуклеотидных микрочипов сдерживается необходимостью применения мягких условий гибридизации, при которых происходит внутреннее спаривание нуклеотидов как в зондах, так и в анализируемых частях генов, что часто является причиной неправильной интерпретации получаемых данных. Тем не менее, на основании результатов гибридизации продуктов ПЦР-амплификации частей клонов геномных библиотек с микроматрицами олигонуклеотидов удается осуществлять достаточно эффективное упорядочивание клонов друг относительно друга. Продемонстрирована возможность осуществления биохимических манипуляций с ДНК, иммобилизованных на микрочипах, с использованием ДНК-полимераз и эндонуклеаз рестрикции. Такие воздействия применяют параллельно с гибридизацией для получения дополнительной информации об анализируемых ДНК.

Исследование генетического полиморфизма ДНК. Более информативными оказываются результаты использования микроматриц ДНК для оценки генетического полиморфизма родственных геномов. Способность зондов гибридизоваться с ДНК микроматриц сильно изменяется при наличии даже единственного неспаренного нуклеотида в гибридах зонд–мишень. Это позволяет с высокой эффективностью осуществлять поиск полиморфизмов в

сравниваемых геномах даже на уровне различий в отдельных нуклеотидах. Например, путем гибридизации геномной ДНК двух штаммов дрожжей с микроматрицами олигонуклеотидов удалось картировать на участке ДНК длиной в 57 т.п.о. полиморфные локусы, детерминирующие множественную лекарственную устойчивость. Детальное сравнение экзона 11 гена BRCA1 человека с гомологичными последовательностями семи других биологических видов позволило обнаружить отдельные полиморфные нуклеотиды с замечательной точностью (99,5–99,9%).

Такого рода результаты дали возможность разработки на основе микроматриц ДНК нескольких ДНК-диагностикумов, с помощью которых обнаруживают полиморфизмы на уровне отдельных нуклеотидов в геноме вирусов, бактерий, дрожжей и высших эукариот. С применением этого подхода были идентифицированы мутации в геноме человека, ассоциированные с различными заболеваниями, включая рак молочной железы, муковисцидоз, β -талассемию, а также диагностирована ВИЧ-инфекция. В недавно проведенном исследовании участка генома человека длиной в $2,3 \cdot 10^6$ п.о. было обнаружено ~2000 полиморфизмов на уровне отдельных нуклеотидов, что позволяет представить себе масштабы генетической изменчивости в популяциях человека.

Исследование экспрессии генов с использованием микроматриц ДНК. Технология микрочипов ДНК позволяет осуществлять одновременный мониторинг за экспрессией большого числа генов (expression profiling). С этой целью для генов с известными последовательностями нуклеотидов создается микроматрица сегментов кДНК длиной 0,5–1,0 т.п.о. Из анализируемых образцов (например опухоли и здоровой ткани) выделяют суммарную мРНК, которую с помощью обратной транскрипции превращают в кДНК, метят флуоресцентными красителями и используют для последующей конкурентной гибридизации с зондами, нанесенными на микроматрицу. Интенсивность флуоресценции отдельных элементов микроматрицы после образования гибридов позволяют качественно характеризовать различия в уровнях экспрессии конкретных генов в анализируемых образцах. Например, отсутствие конкуренции за образование гибридов со стороны кДНК нормальной ткани может говорить о транскрипции в опухоли новых генов, не экспрессирующихся в нормальных клетках.

Такой подход успешно использовали для характеристики ответа клеток в культуре или *in vivo* по изменению уровней экспрессии генов на различные внешние стимулы, включая тепловой шок, воспалительные реакции и канцерогенные воздействия. Помимо известных генов в мониторинг иногда включают и случайные клоны кДНК, что позволяет идентифицировать новые гены, экспрессия которых ассоциирована с патологическими состояниями органов и тканей. С использованием микроматриц кДНК в современный анализ могут быть одновременно включены до 10000 экспрессирующихся генов.

Микроматрицы олигонуклеотидов также применяют для определения профилей экспрессии генов. В этом случае для повышения эффективности мониторинга одновременно в разных элементах матрицы используют несколько олигонуклеотидов, комплементарных различным частям анализируемых кДНК. С помощью этого подхода недавно был определен профиль экспрессии 6800 генов в культивируемых клетках фибросаркомы в присутствии γ -интерферона.

Функциональная геномика. Эффективное использование результатов мониторинга за экспрессией большого числа генов требует лучшего понимания функций анализируемых генов. Решением проблемы определения функций новых генов занимается раздел генетики, получивший название функциональной геномики. Лишь немногие заболевания возникают в результате повреждения отдельных конкретных генов. В большинстве случаев приходится говорить о предрасположенности к заболеванию в связи с наличием в геноме конкретной мутации. Сопоставление генетической информации, получаемой при использовании микроматриц ДНК, с результатами статистического анализа возникновения, протекания и исхода заболеваний может дать ключ к правильной интерпретации результатов генетического скрининга генома человека обсуждаемыми методами. Возможность одновременного наблюдения за изменением экспрессии очень большого числа генов в строго контролируемых условиях открывает новые перспективы функционального исследования генома как единого целого.

В заключение можно сказать, что микроматрицы ДНК уже сегодня находят применение в фундаментальных исследованиях, промышленности и клинике в качестве инструментов анализа экспрессии большого числа генов, определения генетической предрасположенности организма к различным патологическим состояниям, скрининга новых фармакологических препаратов.

Критическими моментами, препятствующими широкому распространению этих технологий, являются ограниченные чувствительность обнаружения гибридизационных сигналов и специфичность гибридизации, трудности в количественной оценке сигналов и обработке большого количества получаемых данных с целью их интерпретации, а также высокая стоимость микрочипов ДНК. Однако микроматрицы ДНК даже в современном виде успешно используются и будут применяться для решения многих задач биологии и медицины.

ГЛАВА 12.КАРТИРОВАНИЕ И ОПРЕДЕЛЕНИЕ ПЕРВИЧНОЙ СТРУКТУРЫ ГЕНОМА ЧЕЛОВЕКА

После краткого рассмотрения основных методов, наиболее часто используемых в молекулярной генетике для исследования структуры и механизмов функционирования генов, представляется целесообразным на примере генома человека подробнее познакомиться с практическим применением этих методов и их модификаций для изучения больших геномов. В целях всестороннего исследования генома человека, этого колоссального по объему хранилища его генетической информации, недавно была разработана и воплощается в жизнь специальная международная программа "Геном человека" ("Human Genome Project"). Основной задачей программы является построение исчерпывающих генетических карт большого разрешения каждой из 24 хромосом человека, которое, в конечном счете, должно завершиться определением полной первичной структуры ДНК этих хромосом. В настоящее время работы по проекту идут полным ходом. В случае успешного его завершения (а это по планам должно произойти в 2003 г.) у человечества появятся перспективы досконального изучения функциональной значимости и механизмов функционирования каждого из его генов, а также генетических механизмов, управляющих биологией человека, и установления причин большинства патологических состояний его организма.

12.1. Основные подходы к картированию генома человека

Решение основной задачи программы "Геном человека" включает три основных этапа. На первом этапе необходимо специфическим образом разделить каждую индивидуальную хромосому на части меньшего размера, позволяющего их дальнейший анализ известными методами. Вторая стадия исследований предполагает определение взаимного расположения этих индивидуальных фрагментов ДНК друг относительно друга и их локализации в самих хромосомах. На завершающем этапе необходимо произвести собственно определение первичной структуры ДНК каждого из охарактеризованных фрагментов хромосом и составить полную непрерывную последовательность их нуклеотидов. Решение задачи не будет полным, если в найденных

последовательностях нуклеотидов не удастся локализовать все гены организма и определить их функциональное значение. Прохождение трех вышеперечисленных этапов требуется не только для получения исчерпывающих характеристик генома человека, но и любого другого генома большого размера.

12.1.1. Генетические карты сцепления

Генетические карты сцепления представляют собой одномерные схемы взаимного расположения *генетических маркеров* на индивидуальных хромосомах. Под генетическими маркерами понимают любые наследуемые фенотипические признаки, различающиеся у отдельных особей. Фенотипические признаки, отвечающие требованиям генетических маркеров, весьма разнообразны. Они включают в себя как особенности поведения или предрасположенность к определенным заболеваниям, так и морфологические признаки целых организмов или их макромолекул, различающихся по структуре. С развитием простых и эффективных методов исследования биологических макромолекул такие признаки, известные под названием *молекулярных маркеров*, стали наиболее часто использоваться при построении генетических карт сцепления. Прежде чем перейти к рассмотрению методов построения таких карт и их значения для исследования генома, необходимо напомнить, что термин "*сцепление*" употребляется в генетике для обозначения вероятности совместной передачи двух признаков от одного из родителей потомству.

При образовании половых клеток (гамет) у животных и растений на стадии мейоза, как правило, происходит синапсис (конъюгация) гомологичных хромосом. Сестринские хроматиды гомологичных хромосом соединяются по всей длине друг с другом, и в результате *кроссинговера* (генетической рекомбинации между хроматидами) происходит обмен их частями. Чем дальше два генетических маркера располагаются друг от друга на хроматиде, тем больше вероятность того, что разрыв хроматиды, необходимый для кроссинговера, произойдет между ними, и два маркера в новой хромосоме, принадлежащей новой гамете, окажутся отделенными друг от друга, т.е. их сцепление нарушится. Единицей сцепления генетических маркеров является *морганида* (единица Моргана, М), которая содержит 100 *сантиморганид* (сМ). 1

сМ соответствует физическому расстоянию на генетической карте между двумя маркерами, рекомбинация между которыми происходит с частотой 1%. Выраженная в парах оснований 1 сМ соответствует ~1 млн п.о. (м.п.о.) ДНК.

Генетические карты сцепления правильно отражают порядок расположения генетических маркеров на хромосомах, однако полученные при этом значения расстояний между ними не соответствуют реальным физическим расстояниям. Обычно данный факт связывают с тем, что эффективность рекомбинации между хроматидами на отдельных участках хромосом может сильно различаться. В частности, она подавлена в гетерохроматиновых участках хромосом. С другой стороны, в хромосомах часто встречаются "горячие точки" рекомбинации. Использование частот рекомбинации для построения физических генетических карт без учета этих факторов будет приводить к искажениям (соответственно занижению или завышению) реальных расстояний между генетическими маркерами. Таким образом, генетические карты сцепления являются наименее точными из всех имеющихся типов генетических карт, и их можно рассматривать только в качестве первого приближения к реальным физическим картам. Тем не менее, на практике именно они и только они позволяют локализовать сложные генетические маркеры (например ассоциированные с симптомами заболевания) на первых этапах исследования и дают возможность их дальнейшего изучения. Необходимо помнить, что в отсутствие кроссинговера все гены, находящиеся на индивидуальной хромосоме, передавались бы от родителей потомству вместе, поскольку они физически сцеплены друг с другом. Поэтому индивидуальные хромосомы образуют *группы сцепления генов*, и одной из первых задач построения генетических карт сцепления является отнесение исследуемого гена или последовательности нуклеотидов к конкретной группе сцепления. В табл. II.4 перечислены современные методы, которые, по данным В.А. МакКьюзика, наиболее часто использовались для построения генетических карт сцепления до конца 1990 г.

Таблица II.4

**Современные методы построения генетических карт
сцепления**

Метод	Число картированных локусов
Гибридизация соматических клеток	1148
Гибридизация in situ	687
Семейный генетический анализ сцепления	466
Определение эффекта дозы	159
Рестрикционное картирование	176
Использование хромосомных aberrаций	123
Использование синтении	110
Сегрегация генов, индуцированная облучением	18
Другие методы	143
Всего	3030

Гибридизация соматических клеток. Одним из наиболее популярных методов отнесения генетического маркера (функционально активного гена) к конкретной группе сцепления является гибридизация (слияние друг с другом) соматических клеток разных биологических видов организмов, один из которых – исследуемый. У межвидовых гибридов соматических клеток в процессе культивирования происходит утрата хромосом преимущественно одного из биологических видов. Потеря хромосом носит, как правило, случайный характер, и образующиеся клоны клеток содержат оставшиеся хромосомы в разных сочетаниях. Анализ клонов, содержащих разные наборы хромосом исследуемого вида, позволяет определить, с какой из этих оставшихся хромосом ассоциирована экспрессия исследуемого маркера, и, следовательно, локализовать ген на конкретной хромосоме.

Гибридизация in situ. Метод гибридизации in situ также широко используется для картирования последовательностей нуклеотидов на хромосомах. С этой целью препараты фиксированных хромосом гибридизуют (инкубируют при повышенной температуре с последующим охлаждением) с исследуемыми последовательностями нуклеотидов, мечеными радиоактивной, флуоресцентной или иной меткой. После отмывания несвязавшейся метки оставшиеся меченые молекулы нуклеиновых кислот оказываются ассоциированными с участками хромосом, содержащими последовательности, комплементарные исследуемым меченым последовательностям нуклеотидов. Полученные гибриды анализируют с помощью микроскопа либо непосредственно, либо после автордиографии. Для этой группы методов характерна более высокая разрешающая способность, чем для гибридизации соматических клеток, поскольку они позволяют локализовать изучаемые последовательности нуклеотидов на хромосомах. По мере выполнения программы "Геном человека" в руках исследователей появляется все больше изолированных последовательностей нуклеотидов, которые можно использовать в качестве зондов для гибридизации in situ. В связи с этим данные методы по частоте использования в последнее время прочно выходят на первое место. Наиболее популярной оказывается группа методов, получивших название *флуоресцентной гибридизации in situ* (fluorescence in situ hybridization – FISH), при проведении которой используются полинуклеотидные зонды, содержащие флуоресцентную метку. В частности, в

1996 г. было опубликовано >600 работ, в которых описано использование этого метода.

Семейный генетический анализ сцепления. Эта группа методов часто используется в медицинской генетике для выявления связи (сцепления) между симптомами заболевания, вызываемого мутацией в неизвестном гене, и другими генетическими маркерами. В данном случае в качестве одного из генетических маркеров выступают сами симптомы заболевания. В геноме человека обнаружено большое количество *полиморфизмов*, в том числе ПДРФ (см. разделы 11.1 и 11.2). ПДРФ распределены более или менее равномерно в геноме человека на расстоянии ~5–10 сМ друг от друга. Чем ближе индивидуальные полиморфные локусы расположены к гену, ответственному за заболевание, тем меньше вероятность их разделения при рекомбинации в мейозе и тем чаще они будут встречаться вместе у больного индивидуума и вместе передаваться от родителей потомству. Клонировав протяженный участок генома, включающий соответствующий полиморфный маркер (его отбор из клонотеки геномной ДНК проводят с помощью зонда), можно одновременно вместе с ним с большой вероятностью выделить ген, вызывающий наследственное заболевание. Такие подходы были, в частности, успешно применены для проведения семейного анализа и выделения соответствующих генов при мышечной дистрофии Дюшенна, кистозном фиброзе почек (муковисцидозе) и миотонической дистрофии. Информативность отдельных ПДРФ генома человека зависит от уровня их гетерозиготности в исследуемой популяции. Мерой информативности ПДРФ как генетического маркера по предложению Д. Ботштейна и соавторов (1980 г.) принято считать значение *содержания полиморфной информации* PIC (polymorphism information content), которое представляет собой отношение числа скрещиваний, в которых хотя бы у одного из родителей исследуемый полиморфный маркер находится в гетерозиготном состоянии, ко всем скрещиваниям.

Определение эффекта дозы гена и использование хромосомных aberrаций. Этими методами обнаруживают корреляции между уровнем экспрессии исследуемого гена и количеством конкретных хромосом в анеуплоидных линиях клеток или структурными перестройками хромосом (хромосомными мутациями – aberrациями). *Анеуплоидией* называют наличие у клетки, ткани или целого организма числа хромосом, не равного типичному для

данного биологического вида. Хромосомные aberrации в виде транслокаций участков хромосом в гетерохроматиновые области тех же самых или других хромосом часто сопровождаются подавлением транскрипции генов, расположенных в транслоцированных участках или в хромосоме-акцепторе (мозаичный эффект положения). Изменение уровня активности исследуемого гена при конкретных транслокациях может быть основанием для картирования его в соответствующей области хромосомы.

Использование синтении. *Синтения* – это структурное сходство групп сцепления генов у организмов разных биологических видов. В частности, в геномах человека и мыши известно несколько десятков синтеничных групп генов. Наличие феномена синтении позволяет суживать круг поиска места локализации исследуемого гена на хромосомах, ограничивая его областью известных генов, принадлежащих к конкретной синтеничной группе.

Сегрегация генов, индуцируемая ионизирующим излучением. С помощью этого метода определяют расстояние между исследуемыми генами путем оценки вероятности их разделения (сегрегации) после облучения клеток определенной стандартной дозой ионизирующего излучения. Облученные клетки спасают от гибели гибридизацией с соматическими клетками грызунов, и у соматических гибридов в культуре определяют наличие исследуемых маркеров облученных клеток. В итоге удастся сделать вывод о наличии или отсутствии сцепления (физическом расстоянии) между этими генами.

Среди **других методов** следует упомянуть способы, основанные на использовании для картирования генов больших фрагментов ДНК, образуемых под действием крупнощепящих рестриктаз. После расщепления геномной ДНК образующиеся фрагменты разделяют электрофорезом в импульсном электрическом поле и далее их гибридизуют по Саузерну с зондами, соответствующими картируемому гену. Если после проведения гибридизации сигналы обоих зондов локализуются на одном и том же крупном фрагменте ДНК, это говорит о тесном сцеплении таких генов.

12.1.2. ПЦР в исследованиях генома человека

Полимеразная цепная реакция занимает центральное место в разработке подходов к практическому осуществлению программы "Геном человека". Как уже обсуждалось выше (раздел 7.1.3), с помощью ПЦР можно

быстро и эффективно амплифицировать почти любой короткий участок генома человека, и полученные продукты ПЦР далее использовать в качестве зондов для картирования соответствующих участков на хромосомах путем гибридизации по Саузерну или *in situ*.

Концепция STS. Одной из ключевых концепций, лежащих в основе картирования генов человека в рамках обсуждаемой программы, является концепция *сайтов, привязанных к последовательностям* (sequence-tagged sites – STS). В соответствии с этой концепцией все фрагменты ДНК, используемые для построения генетических или физических карт, можно однозначно идентифицировать с помощью последовательности нуклеотидов длиной в 200–500 п.о., которая будет уникальной для данного фрагмента. Каждый из этих сайтов необходимо секвенировать, что даст возможность в дальнейшем их амплифицировать с помощью ПЦР и применять в качестве зондов. Использование STS позволило бы применять их последовательности в виде продуктов ПЦР в качестве зондов для направленного выделения любого фрагмента ДНК того или иного участка генома из клонотек геномных последовательностей. В результате могут быть созданы базы данных, включающие локализацию и структуру всех STS, а также праймеров, необходимых для их амплификации. Это избавило бы лаборатории от необходимости хранения многочисленных клонов и их рассылки в другие лаборатории для проведения исследований. Кроме того, STS создают основу для разработки единого языка, на котором разные лаборатории могли бы описывать свои клоны. Таким образом, конечным результатом разработки концепции STS была бы исчерпывающая карта STS генома человека. Теоретически для построения генетической карты размером в 1 сМ необходимо ~3000 полностью информативных, полиморфных ДНК-маркеров. Однако поскольку полиморфные маркеры распределены в геноме неравномерно и лишь немногие из них полностью информативны, реальное число маркеров, требуемых для построения карты такого размера, оценивается в 30–50 тысяч. Для получения маркеров, соответствующих исследуемым участкам хромосом, в настоящее время часто применяют праймеры, соответствующие диспергированным повторяющимся последовательностям, среди которых первыми стали использовать *Alu*-последовательности (см. раздел 1.3.1).

***Alu*-ПЦР.** Диспергированные повторяющиеся *Alu*-последовательности

характерны именно для генома человека. Праймеры, специфичные в отношении *Alu*-последовательностей, используют для амплификации участков ДНК генома человека, заключенных между *Alu*-повторами, которые располагаются в среднем на расстоянии 4–10 т.п.о. друг от друга. Другим вариантом *Alu*-ПЦР является направленный синтез с ее помощью ДНК-зондов к участкам хромосом, полученным после лазерной фрагментации, индивидуальным хромосомам, выделенным с помощью проточной цитофлуориметрии, или ДНК гибридных клеток, содержащих определенную часть генома человека. Кроме того, *Alu*-ПЦР используют для получения уникальных фингерпринтов, характеризующих клеточные гибриды с точки зрения стабильности их генома, а также для характеристики фрагментов ДНК человека, клонированных в YAC-векторах, космидах или векторах на основе ДНК бактериофагов. Уникальность *Alu*-последовательностей для генома человека делает возможным их применение для "прогулок по хромосомам" (см. раздел 7.4.2), а также для расширения существующих *контигов* (см. ниже). Поскольку в геноме человека >90% умеренно повторяющихся последовательностей представлены семействами *Alu* и *KpnI*, неудивительно, что последние также применяются в ПЦР для тех же целей, что и *Alu*. Однако здесь профили продуктов ПЦР менее сложны, поскольку последовательности *KpnI* повторяются в геноме реже и обладают характерной локализацией в хромосомах.

ПЦР активно используется для выявления полиморфных молекулярных маркеров при построении генетических карт сцепления, основные принципы получения которых были рассмотрены выше. Этот метод оказывается полезным и при секвенировании ДНК, а также при построении физических карт высокого разрешения для генома человека. О последних двух сферах применения ПЦР подробнее речь пойдет ниже.

12.1.3. Физические карты низкого разрешения

В отличие от рассмотренных выше генетических карт сцепления физические карты генома отражают реальное расстояние между маркерами, выражаемое в парах оснований. Физические карты различаются по степени их разрешения, т.е. по тем деталям структуры генома, которые на них представлены (рис. II.39). Исчерпывающая физическая карта генома человека

максимального разрешения будет содержать полную нуклеотидную последовательность всех его хромосом. На другом полюсе физических карт с минимальным разрешением находятся *хромосомные (цитогенетические) карты* генома.

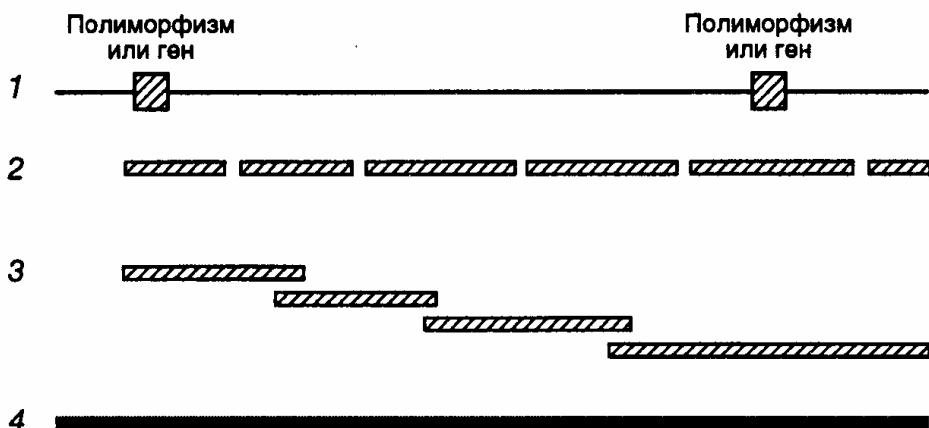


Рис. II.39. Четыре типа генетических карт геномной ДНК и их взаимоотношения

1 – генетическая карта сцепления, 2 – физическая рестрикционная карта, пробелы обозначают места расщепления ДНК рестриктазами, 3 – физическая карта контигов, показаны перекрывающиеся клоны ДНК, полученные с помощью YAC-векторов, 4 – исчерпывающая физическая карта в виде последовательности нуклеотидов ДНК. На всех картах представлен один и тот же участок хромосомы

Хромосомные карты. Хромосомные карты генома человека получают локализацией генетических маркеров на индивидуальных хромосомах с использованием цитогенетических методов, включая автордиографию и FISH. В последних двух случаях радиоактивная или флуоресцентная метки, ассоциированные с исследуемыми генетическими локусами интактных хромосом, выявляются с помощью световой микроскопии. Еще совсем недавно хромосомные карты позволяли локализовать исследуемый фрагмент ДНК на участке хромосомы протяженностью ~10 м.п.о. Современные методы гибридизации *in situ* с использованием метафазных хромосом, главным образом, метод FISH, локализуют полинуклеотидные маркеры в пределах 2–5 м.п.о. Более того, при гибридизации *in situ* с интерфазными хромосомами, в которых генетический материал находится в менее компактной форме, разрешающая способность хромосомных карт приближается к 100 т.п.о.

Точность хромосомных карт повышается и с использованием

современных генетических методов. Например, способность ПЦР амплифицировать сегменты ДНК единичного сперматозоида позволяет исследовать большое число мейозов, как бы законсервированных в отдельных образцах спермы. В результате появляется возможность проверки взаимного расположения генетических маркеров, локализованных на хромосомных картах более грубыми методами.

Карты кДНК. Карты кДНК отражают положение экспрессирующихся участков ДНК (экзонов) относительно известных цитогенетических маркеров (бэндов) на метафазных хромосомах. Поскольку такие карты дают представление о локализации транскрибирующихся участков генома, в том числе и генов с неизвестными функциями, они могут быть использованы для поиска новых генов. Этот подход особенно полезен при поиске генов, повреждения которых вызывают заболевания человека, в том случае если приблизительная локализация таких участков хромосом уже предварительно проведена на генетических картах сцепления в результате семейного генетического анализа.

12.1.4.

Физические карты высокого разрешения

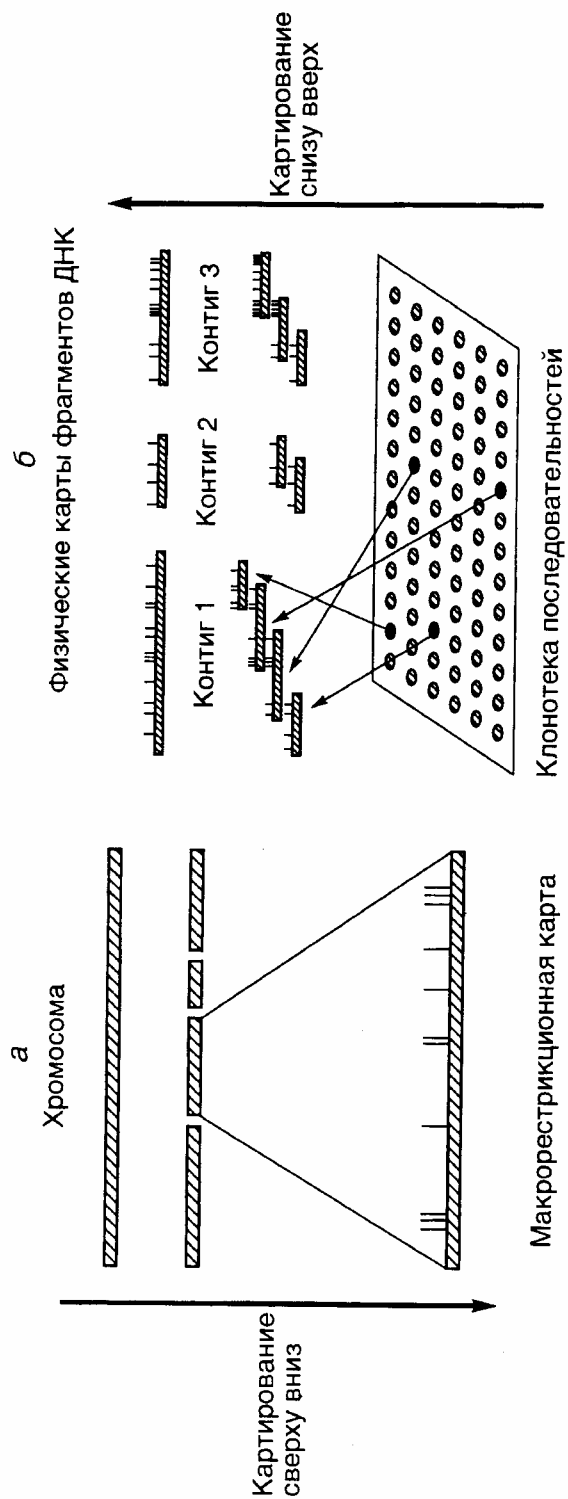


Рис. II.40. Две стратегии построения физических карт ДНК

а – стратегия "сверху вниз": ДНК целой хромосомы расщепляется крупнощепящими рестриктазами, для каждого из индивидуальных фрагментов ДНК строится рестрикционная карта; б – стратегия "снизу вверх", индивидуальные YAC-клоны после идентификации объединяются в контиги

В попытках построения карт генома человека высокого разрешения экспериментально реализуются два альтернативных подхода, получивших названия *картирования сверху вниз* (top-down mapping) и *картирования снизу вверх* (bottom-up mapping). При картировании сверху вниз (рис. II.40,а) исходным в анализе является препарат ДНК индивидуальной хромосомы человека. ДНК разрезается крупнощепящими рестриктазами (например *NotI*) на длинные фрагменты, которые после разделения электрофорезом в импульсном электрическом поле подвергаются дальнейшему рестрикционному анализу с другими рестриктазами. В результате получают *макрорестрикционную карту*, на которой достаточно полно представлены все последовательности исследуемой хромосомы или ее части, однако ее разрешение невысоко. На такой карте очень трудно локализовать индивидуальные гены. К тому же каждая индивидуальная карта редко охватывает протяженные сегменты ДНК (как правило, не более 1–10 м.п.о.).

При картировании генома человека снизу вверх (см. рис. II.40,б) на основе препарата суммарной ДНК генома или индивидуальной хромосомы получают серию случайных клонов протяженных последовательностей ДНК (10–1000 т.п.о), часть из которых перекрывается друг с другом. В качестве вектора для клонирования в этом случае часто используют искусственные минихромосомы бактерий (BAC) или дрожжей (YAC), подробно описанные в разделе 7.2.4. Серия частично перекрывающихся и дополняющих друг друга клонов образует непрерывную состыкованную (contiguous) последовательность нуклеотидов ДНК, получившую название контига (contig). Правильность полученных контигов подтверждают гибридизацией *in situ* (FISH) с одновременной их привязкой к определенным участкам исследуемых хромосом. Карты, основанные на контигах, представляют полную информацию о структуре отдельных сегментов хромосом и позволяют локализовать отдельные гены. Однако такие карты трудно применять для реконструкции целых хромосом или протяженных их участков из-за отсутствия

соответствующих клонов в имеющихся клонотеках генов.

Основная проблема, которую приходится решать при использовании обоих подходов к построению физических карт высокого разрешения, – объединение разрозненных фрагментов ДНК в непрерывные последовательности нуклеотидов. Чаще всего для этого применяют специальные клонированные фрагменты ДНК, получившие название *связующих* (linking) *клонов*. Фрагменты ДНК из связующих клонов содержат в своих внутренних частях последовательности нуклеотидов крупнощепящих рестриктаз и, следовательно, представляют собой места стыковки фрагментов ДНК, используемых на первых этапах физического картирования. Гибридизацией по Саузерну, при проведении которой в качестве зондов используют фрагменты ДНК связующих клонов, определяют фрагменты ДНК физических карт, содержащие последовательности нуклеотидов окрестностей сайтов рестрикции крупнощепящих рестриктаз. Если два таких фрагмента найдены, то соответствующий связующий клон перекрывает оба этих фрагмента и является их частью. Связующие клоны, в свою очередь, отбирают из клонотек генов с помощью зондов, которые представляют собой последовательности нуклеотидов сайтов рестрикции крупнощепящих рестриктаз.

12.2. Определение полной первичной структуры ДНК генома человека

Исчерпывающая физическая карта генома человека (и любого другого организма) должна представлять собой полную последовательность нуклеотидов ДНК всех его хромосом. Благодаря тому, что к решению такой грандиозной по объему задачи привлечены многие хорошо финансируемые лаборатории в разных странах мира и для этих целей используются автоматические высокопроизводительные секвенаторы, она представляется скорее рутинной, чем захватывающей. Секвенирование начинается с субклонирования небольших рестрикционных фрагментов ДНК, образуемых в результате рестрикции предварительно клонированных более крупных фрагментов, и первичная структура коротких фрагментов далее определяется одним из известных методов.

Опубликованные одновременно в 1977 г. методы секвенирования ДНК А. Максама и У. Гилберта, а также Ф. Сенгера сделали процесс определения

первичной структуры генов доступным любой биохимической лаборатории. С тех пор метод Сенгера, основанный на случайной терминации синтеза ДНК *in vitro* как следствие включающихся в синтезируемую цепь ДНК дидезоксирибонуклеозидмонофосфатов, получил широкое распространение. Дальнейшим развитием метода Сенгера было использование в продукте реакции флуоресцентной метки вместо радиоактивной, что позволило автоматизировать процесс секвенирования. Во многих случаях для определения первичной структуры того или иного участка ДНК теперь уже нет необходимости в предварительном клонировании соответствующего фрагмента, поскольку разработаны эффективные методы секвенирования непосредственно продуктов ПЦР. Все это резко стимулировало эффективность анализа первичной структуры генов. Однако необходимо иметь в виду, что длина ДНК даже самой маленькой Y-хромосомы человека составляет 50 м.п.о., а лучшие автоматические секвенаторы сегодняшнего дня позволяют определять за один год последовательность из ~0,1 м.п.о. Это наглядно демонстрирует масштабы задачи, которую предстоит решить³.

12.3. Базы данных получаемой информации

Полное использование информации о структуре генома человека в биологии и медицине станет возможным лишь в отдаленном будущем. Еще долгие годы предстоит собирать и обрабатывать получаемую информацию, устранять противоречия между экспериментальными данными разных лабораторий. Одна только запись всей последовательности нуклеотидов генома человека обычным шрифтом потребует ~200 томов книг по 1000 страниц каждая. Анализ такого объема информации неминуемо повлечет за собой необходимость разработки новых подходов в информатике и будет невозможен без создания уникальных алгоритмов обработки информации с помощью вычислительной техники.

По мере накопления информации о последовательностях нуклеотидов различных частей генома человека на первое место будут выходить задачи по идентификации среди них функционально значимых. В первой части книги уже

³ К концу 1999 года полностью определена последовательность нуклеотидов хромосомы 22 человека.

подчеркивалось, что даже определение границ генов в ряде случаев является трудно разрешимой задачей. То же относится и к идентификации регуляторных последовательностей генов. Не исключено, что накопление новых данных о структуре генома человека, в конце концов, потребует пересмотра как критериев функциональной значимости нуклеотидных последовательностей генома, так и самой концепции гена.

Крупнейшей базой данных по структуре генома человека в настоящее время является GDB (Genome Data Base), созданная и поддерживаемая в университете Дж. Гопкинса (Балтимор, США). Кроме известных последовательностей нуклеотидов генома человека в ней хранится вся получаемая информация о генетических маркерах, зондах и контигах, ассоциированных с генетическими заболеваниями. Проводятся работы по включению в базу данных результатов физического картирования генома. В этом же университете поддерживается база данных по менделевскому наследованию у человека (Online Mendelian Inheritance in Man Database), которая представляет собой каталог наследуемых признаков и наследственных заболеваний человека.

Четыре другие базы данных хранят все известные последовательности нуклеотидов, включая последовательности нуклеотидов генома человека: GenBank и Genome Sequence Data Base (GSDB) в США, European Molecular Biology Laboratory (EMBL) Nucleotide Sequence Database в Великобритании, а также DNA Data Bank of Japan (DDBJ). В этих базах данных в 1996 году хранилось ~200 м.п.о. последовательностей нуклеотидов, предоставленных как самими авторами, так и вводимых из статей, опубликованных в периодических изданиях. В России Институтом молекулярной биологии РАН поддерживается аналогичная база данных по геному человека (Hugene) с более скромными задачами, работающая на отечественном программном обеспечении. Основной международной базой данных по последовательностям аминокислот является Protein Identification Resource (Швейцария).

Принесет ли молекулярной генетике определение полной первичной структуры генома человека принципиально новые знания? Действительно, максимальную информацию о функционировании генома и функциональной значимости его основных частей могло бы принести только полное сравнительное секвенирование двух геномов, принадлежащих разным

индивидуумам. Это позволило бы отделить консервативные последовательности от последовательностей, изменяющихся в процессе онтогенеза. Ведь даже в двух соседних соматических клетках одного организма последовательности нуклеотидов различаются из-за соматического мутагенеза. Выполняемая же программа исследований по геному человека в конечном счете даст только усредненную информацию. Если же, как это предполагается в разделе 5.3, основной функцией избыточных последовательностей, составляющих ~90% генома млекопитающих, является обеспечение запрограммированной скорости изменения отдельных генетических локусов под действием мутаций, и(или) поддержания пространственной структуры кодирующих частей генов, то полное секвенирование таких последовательностей будет лишено большого смысла. Каков будет конец этой истории, может показать только будущее.

ЗАКЛЮЧЕНИЕ

Современная генетика находится на взлете. Новые факты обнаруживаются настолько быстро, что едва хватает времени на то, чтобы просто осознать их появление. Еще труднее уловить многочисленные связи между ними. Уже сейчас многие главы монографии хотелось бы переписать заново и реализацию этого замысла сдерживает лишь бесконечность такого пути. В этом заключении я хотел бы, с известной долей субъективности, отметить несколько новых направлений генетических исследований, которые достаточно явно обозначились к настоящему времени и, на мой взгляд, могут определять некоторые будущие направления генетических исследований.

Геномика. Конец 1990-х годов отмечен определением полной первичной структуры геномов нескольких организмов. В настоящее время уже имеются полные данные о последовательности нуклеотидов целых геномов нескольких зубактерий, архебактерии и эукариотического организма – дрожжей. На подходе получение таких данных о геномах дрозофилы, нематоды *C. elegans*, арабидопсиса и человека. По мере развития этих исследований молекулярная биология получит через соответствующие базы данных доступ к неограниченному числу первичных структур природных белков и нуклеиновых кислот. Ключевой проблемой, которую необходимо решить в связи с накоплением такого рода информации (и которая пока далека от своего разрешения) является возможность соотнесения первичных структур открываемых новых генов с функциями кодируемых этими генами белков и нуклеиновых кислот.

Базы данных последовательностей нуклеотидов целых геномов дают возможность изучения эволюционной истории семейств белков разных видов организмов. На основании этих данных все аминокислотные последовательности организуют в виде набора ~10 000 независимо эволюционирующих блоков ("модулей"). Для каждого из этих модулей на основании степени сходства или различия их аминокислотных (и соответствующих нуклеотидных) последовательностей строится эволюционное древо, которое отражает эволюционную историю семейств белков и позволяет

обнаруживать предполагаемые исходные полипептиды-предшественники. При этом полагают (С.Ф. Беннер с соавторами, 1998 г.), что сопоставление данных эволюционной истории белков с их биохимическими свойствами и пространственной структурой позволит, в конечном счете, связать первичную структуру полипептидов с их биологическими функциями. Решение этой ключевой проблемы геномики будет сопровождаться получением большого количества новой информации о конформации макромолекул, их надмолекулярной организации, механизмах катализа и несомненно обогатит многие области исследований в биологии и химии.

Экспрессией генов управляют большие надмолекулярные комплексы. Определение функций продуктов новых генов на основе их первичной структуры осложняется тем, что многие белки и нуклеиновые кислоты проявляют свою активность и функционируют только в составе больших надмолекулярных комплексов, размеры которых часто приближаются к размерам рибосом. При этом многие белки сами по себе не обладают ферментативной активностью, а выполняют вспомогательные функции, например, молекул-адаптеров, обеспечивающих сборку комплексов и создающих молекулярные интерфейсы для их взаимодействия с регуляторными и каталитическими субъединицами. Наличие таких комплексов, как это было установлено в последнее десятилетие, особенно характерно для клеток эукариотических организмов. Так, исследование молекулярных механизмов транскрипции у эукариот привело к развитию представления о транскриптосоме, гигантскому белковому комплексу, в который кроме холофермента РНК-полимеразы с ее многочисленными субъединицами входят факторы транскрипции, белки-адаптеры, белковые компоненты системы репарации и т.п. При этом размер транскриптосомы приближается к таковому целых рибосом. В гигантские надмолекулярные комплексы организованы и молекулярные машины системы синтеза ДНК (реплисомы), процессинга и редактирования РНК (сплайсомы и эдитосомы), молекулярные компоненты системы протеолитической деградации белков (протеасомы). Создается впечатление, что организация генетических систем, функционирующих на основных этапах реализации генетической информации, в гигантские пространственно упорядоченные комплексы, является общебиологическим принципом.

Комбинаторные принципы. Важным достижением исследований последних лет является обнаружение способности объединения факторов транскрипции *в разных сочетаниях* в большие белковые комплексы. При этом каждое новое сочетание ограниченного числа факторов придает комплексу уникальные регуляторные свойства. Возможность такого объединения определяется наличием в их полипептидных цепях гомологичных доменов, которые обеспечивают соответствующие белок-белковые взаимодействия. Реализация комбинаторного принципа позволяет клетке расширить регуляторные возможности системы транскрипции без привлечения новых генов, то есть более эффективно использовать генетическую информацию своего генома.

Принципы комбинаторики стали находить широкое применение в прикладных молекулярно-биологических исследованиях. Использование автоматических синтезаторов нуклеиновых кислот позволяет с легкостью получать в одной пробирке наборы олигонуклеотидов, содержащие *все теоретически возможные* нуклеотидные последовательности и следовательно ассоциированные с ними (прямо или косвенно) возможные биологические активности. Разработанные простые системы скрининга позволяют выделять из пула случайных последовательностей нуклеотидов аптамеры и молекулы рибозимов, обладающие требуемыми биологическими свойствами. С помощью систем, аналогичных фаговому дисплею, среди продуктов трансляции таких олигонуклеотидов обнаруживают новые биологически-активные пептиды.

Четырехмерные модели организации экспрессии генов. Современные карты метаболических путей, с помощью которых пытаются с исчерпывающей полнотой представить все последовательности биохимических реакций, протекающих в клетке, своей чрезмерной сложностью обнаруживают слабость данного прямолинейного подхода к отображению биохимической информации. Попытки разворачивания на плоскости в виде двумерной карты всей последовательности биохимических реакций, высокоупорядоченных в пространстве и во времени, делают такие всеобъемлющие схемы малоприменимыми на практике. С аналогичными трудностями придется столкнуться и генетикам при построении моделей, которые бы отображали многочисленные отношения между экспрессирующимися генами. Делом будущего является установление физических и функциональных связей между

надмолекулярными комплексами высокоорганизованных в пространстве генетических подсистем, которые упорядоченно изменяются во времени. Реализация такого комплексного подхода к анализу генетических систем позволит отчетливо увидеть в цитозоле и органеллах клеток единую генетическую систему организма.

Генетический индетерминизм: генотип не определяет всей нормы реакции организма. Программа-максимум, которую можно было бы сформулировать, исходя из будущих достижений геномики, это предсказание всех морфологических и физиологических особенностей взрослого организма на основании одной лишь первичной структуры его генома или генотипа его родителей. Возможно ли это? Такая постановка вопроса напоминает рассуждения философов-детерминистов о предсказании любого события в мире при наличии *полных* знаний об исходных условиях, в которых оно происходит, а также о причинно-следственных связях между явлениями. История науки решила эту проблему не в пользу детерминистов. Выбор путей перехода сложной физической системы в альтернативные состояния часто бывает случайным. И в этом отношении живой организм, по-видимому, не является исключением.

Неопределенность преобразования генома в индивидуальном развитии.

Неопределенность понятия гена.

Случайность и необходимость генетических изменений.

Случайность и необходимость в экспрессии генов.

Конечный результат экспрессии генов предопределен.

Будущее трансгеноза и генотерапии. Это будет. И совершенно безразлично - хотим мы этого или нет.

Большинство физиологических моделей, в которых делается попытка описания регуляторных воздействий на промежуточные или конечные результаты экспрессии генов на клеточном или организменном уровне, носят качественный характер. Это определяется необычайной сложностью даже самых простых биологических (а следовательно и генетических) объектов исследования. Такое состояние дел не может удовлетворить нас,

воспитанников школы физико-химической биологии.

Регуляция экспрессии генов: информационные сети

Все меньше надежды на простоту, подвижный мобиль

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

К введению

Barbieri M. The organic codes. The basic mechanism of macroevolution. // *Biology Forum*. 1998. Vol. 91. P. 481–514.

К главе 1

Георгиев Г.П. Гены высших организмов и их экспрессия. М.: Наука, 1989. 254 с.

Георгиев Г.П., Бакаев В.В. Три уровня структурной организации хромосом эукариот // *Молекуляр. биология*. 1978. Т.12. С. 1205–1230.

Глазков М.В. Петельно-доменная организация генов в эукариотических хромосомах // Там же. 1995. Т.29. С. 965–982.

Жимулев И.Ф. Хромомерная организация политенных хромосом. Новосибирск: Наука, 1993. 564 с.

Молекулярная биология. Структура и биосинтез нуклеиновых кислот // Под ред. А.С. Спирина. М.: Высш. шк. 1990. 352 с.

Adams C.R., Kamakaka R.T. Chromatin assembly: Biochemical identities and genetic redundancy // *Curr. Opinion Genet. Develop.* 1999. Vol. 9. 185–190.

Debrauwere H., Gendrel C.G., Lechat S., Dutreix M. Differences and similarities between various tandem repeat sequences: Minisatellites and microsatellites // *Biochimie*. 1997. Vol. 79. P. 577–586.

Gregory T.R., Hebert P.D.N. The modulation of DNA content: Proximate causes and ultimate consequences // *Genome Res.* 1999. Vol. 9. P. 317–324.

Gruss C., Knippers R. Structure of replicating chromatin // *Progr. Nucl. Acids Res. Mol. Biol.* 1996. Vol. 52. P. 337–365.

Hart C.M., Laemmli U.K. Facilitation of chromatin dynamics by SARs // *Curr. Opinion Genet. Develop.* 1998. Vol. 8. P. 519–525.

- van Holde K., Zlatanova J.* What determines the folding of the chromatin fiber? // Proc. Natl. Acad. Sci. USA. 1996. Vol. 93. P. 10548–10555.
- Kornberg R.D., Lorch Y.* Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome // Cell 1999. Vol. 98. P. 285–294.
- Mushegian A.R., Koonin E.V.* A minimal gene set for cellular life derived by comparison of complete bacterial genomes // Proc. Nat. Acad. Sci. US. 1996. Vol. 93. P. 10268–10273.
- Olsen G.J., Woese C.R.* Lessons from an Archaeal genome: What are we learning from *Methanococcus jannaschii*? // Trends Genet. 1996. Vol. 12. P. 377–379.
- Poljak L., Kas E.* Resolving the role of topoisomerase II in chromatin structure and function // Trends Cell Biol. 1995. Vol. 5. P. 348–354.
- Razin S.V.* Functional architecture of chromosomal DNA domains // Crit. Rev. Eukaryotic Gene Expression. 1996. Vol. 6. P. 247–269.
- Robinow C., Kellenberger E.* The bacterial nucleoid revisited // Microbiol. Rev. 1994. Vol. 58. P. 211–232.
- Workman J.L., Kingston R.E.* Alteration of nucleosome structure as a mechanism of transcriptional regulation // Annu. Rev. Biochem. 1998. Vol. 67. PP. 545–579.

К главе 2

- Молекулярная биология: Структура и биосинтез нуклеиновых кислот / Под ред. А.С. Спирина. М.: Высш. шк. 1990. 352 с.
- Спирин А.С.* Молекулярная биология. Структура рибосомы и биосинтез белка. М.: Высш. шк., 1986. 303 с.
- Chakraborty K.* Functional interaction of yeast elongation factor 3 with yeast ribosomes // Intern. J. Biochem. Cell Biol. 1999. Vol. 31. P. 163–173.
- Cate J.H., Yusupov M.M., Yusupova G.Zh. et al.* X-ray crystal structures of 70S ribosome functional complexes // Science. 1999. Vol. 285. P. 2095–2104.
- Edmondson D.G., Roth S.Y.* Chromatin and transcription // FASEB J. 1996. Vol. 10. P. 1173–1182.

- Frank J.* The ribosome at higher resolution – the donut takes shape // *Curr. Opinion Struct. Biol.* 1997. Vol. 7. P. 266–272.
- Futterer J., Hohn T.* Translation in plants – rules and exceptions // *Plant Mol. Biol.* 1996. Vol. 32. P. 159–189.
- Green R., Noller H.F.* Ribosomes and translation // *Annu. Rev. Biochem.* 1997. Vol. 66. P. 679–716.
- Gregory P.D., Hörz W.* Life with nucleosomes: Chromatin remodeling in gene regulation // *Curr. Opinion Cell Biol.* 1998. Vol. 10. P. 339–345.
- Hampsey M.* Molecular genetics of the RNA polymerase II general transcriptional machinery // *Microbiol. Mol. Biol. Rev.* 1998. Vol. 62. P. 465–503.
- Hodges P., Scott J.* Apolipoprotein B mRNA editing: A new tier for control of gene expression // *Trends Biochem. Sci.* 1992. Vol. 17. P. 77–81.
- Koleske A.J., Young R.A.* The RNA polymerase holoenzyme and its implications for gene regulation // *Ibid.* 1995. Vol. 20. P. 113–116.
- Lamond A.I., Earnshaw W.C.* Structure and function in the nucleus // *Science.* 1998. Vol. 280. P. 547–553.
- Latchman D.S.* Transcription factors: An overview // *Intern. J. Biochem. Cell. Biol.* 1997. Vol. 29. P. 1305–1312.
- Lewis J.D., Izaurralde E.* The role of the cap structure in RNA processing and nuclear export // *Europ. J. Biochem.* 1997. Vol. 247. P. 461–469.
- McCarthy J.E.G.* Posttranscriptional control of gene expression in yeast // *Microbiol. Mol. Biol. Rev.* 1998. Vol. 62. P. 1492–1553.
- McKendrick L., Pain V.M., Morley S.J.* Translation initiation factor 4E // *Intern. J. Biochem. Cell. Biol.* 1999. Vol. 31. P. 31–35.
- Morse R.H.* Transcribed chromatin // *Trends Biochem. Sci.* 1993. Vol. 17. P. 23–25.
- Orphanides G., Lagrange T., Reinberg D.* The general transcription factors of RNA polymerase II // *Gen. Develop.* 1996. Vol. 10. P. 2657–2683.
- Reines D., Conaway J.W., Conaway R.C.* The RNA polymerase II general elongation factors // *Trends Biochem. Sci.* 1996. Vol. 21. P. 351–355.

- Smith H.C., Sowden M.P.* Base-modification mRNA editing through deamination – the good, the bad and the unregulated // *Trends Genet.* 1996. Vol. 12. P. 418–424.
- Sugiura M., Hirose T., Sugita M.* Evolution and mechanism of translation in chloroplasts // *Annu. Rev. Genet.* 1998. Vol. 32. P. 437–459.
- Taanman J.-W.* The mitochondrial genome: Structure, transcription, translation and replication // *Biochim. et Biophys. acta.* 1999. Vol. 1410. P. 103–123.
- Tsukiyama T., Wu C.* Chromatin remodeling and transcription // *Curr. Opinion Genet. Develop.* 1997. Vol. 7. P. 182–191.
- Uptain S.M., Kane C.M., Chamberlin M.J.* Basic mechanisms of transcript elongation and its regulation // *Annu. Rev. Biochem.* 1997. Vol. 66. P. 117–172.
- Wittmann H.-G.* Ribosomen und Proteinbiosynthese // *Biol. Chem. Hoppe-Seyler.* 1989. Vol. 370. P. 87–99.
- Workman J.L., Kingston R.E.* Alteration of nucleosome structure as a mechanism of transcriptional regulation // *Annu. Rev. Biochem.* 1998. Vol. 67. P. 545–579.
- Yarnell W.S., Roberts J.W.* Mechanism of intrinsic transcription termination and antitermination // *Science.* 1999. Vol. 284. P. 611–615.

К главе 3

- Жимулев И.Ф.* Гетерохроматин и эффект положения гена. Новосибирск: Наука, 1993. 491 с.
- Льюин Б.* Гены. М.: Мир, 1987. 544 с.
- Молекулярная биология: Структура и биосинтез нуклеиновых кислот / Под ред. А.С. Спирина. М. Высш. шк., 1990. 352 с.
- Bashirullah A., Cooperstock R.L., Lipshitz H.D.* RNA localization in development // *Annu. Rev. Biochem.* 1998. Vol. 67. P. 335–394.
- Blumenthal T.* Trans-splicing and polycistronic transcription in *Caenorhabditis elegans* // *Trends Genet.* 1995. Vol. 11. P. 132–136.
- Bock A., Forchhammer K., Heider J. et al.* Selenoprotein synthesis: An expansion of

- the genetic code. // Trends Biochem. Sci. 1991. Vol. 16. P. 463–467.
- Bochtler M., Ditzel L., Groll M. et al.* The proteasome // Annu. Rev. Biophys. Biomol. Struct. 1999. Vol. 28. P. 295–317.
- Bohley P.* The fate of proteins in cells // Naturwissenschaften. 1995. Vol. 82. P. 544–550.
- Calkoven C.F., Geert A.B.* Multiple steps in the regulation of transcription-factor level and activity // Biochem. J. 1996. Vol. 317. P. 329–342.
- Callis J.* Regulation of protein degradation // The Plant Cell. 1995. Vol. 7. P. 845–857.
- Chabot B.* Directing alternative splicing: Cast and scenarios // Trends Genet. 1996. Vol. 12. P. 472–478.
- Clark A.L., Docherty K.* Negative regulation of transcription in eukaryotes // Biochem.J. 1993. Vol. 295. P. 521–541.
- Cooper A.A., Stevens T.H.* Protein splicing, self-splicing of genetically mobile elements at the protein level // Trends Biochem. Sci. 1995. Vol. 20. P. 351–356.
- Evdokimova V.M., Ovchinnikov L.P.* Translational regulation by Y-box transcription factor: Involvement of the major mRNA-associated protein, p50 // Intern. J. Biochem. Cell Biol. 1999. Vol. 31. P. 139–149.
- Farabaugh P.J.* Programmed translational frameshifting // Microbiol. Rev. 1996. Vol. 60. P. 104–134.
- Futterer J., Hohn T.* Translation in plants – rules and exceptions // Plant Mol. Biol. 1996. Vol. 32. P. 159–189.
- Gerasimova T.I., Corces V.G.* Boundary and insulator elements in chromosomes // Curr. Opinion Genet. Develop. 1996. Vol. 6. P. 185–192.
- Guarente L.* Transcriptional coactivators in yeast and beyond // Trends Biochem. Sci. 1995. Vol. 20. P. 517–521.
- Jentsch S., Schlenker S.* Selective protein degradation: A journey's end within the proteasome // Cell. 1995. Vol. 82. P. 881–884.
- Latchman D.S.* Eukaryotic transcription factors // Biochem J. 1990. Vol. 270. P. 281–289.

- Lyon M.F.* X-chromosome inactivation // *Curr. Biol.* 1999. Vol. 9. P. R235–R237.
- McCarthy J.E.G.* Posttranscriptional control of gene expression in yeast // *Microbiol. Mol. Biol. Rev.* 1998. Vol. 62. P. 1492–1553.
- Mor A., Amiche M., Nicolas P.* Enter a new post-translational modification: D-amino acids in gene-encoded peptides // *Trends Biochem. Sci.* 1992. Vol. 17. P. 481–485.
- Morse R.H.* Transcribed chromatin // *Trends Biochem. Sci.* 1992. Vol. 17. P. 23–26.
- Orphanides G., Lagrange T., Reinberg D.* The general transcription factors of RNA polymerase II // *Genes Develop.* 1996. Vol. 10. P. 2657–2683.
- Raught B., Gingras A.-C.* eIF4E activity is regulated at multiple levels // *Intern. J. Biochem. Cell Biol.* 1999. Vol. 31. P. 43–57.
- Struhl K.* Fundamentally different logic of gene regulation in eukaryotes and prokaryotes // *Cell.* 1999. Vol. 98. P. 1–4.
- Udvary A.* Dividing the empire: Boundary chromatin elements delimit the territory of enhancers // *EMBO J.* 1999. Vol. 18. P. 1–8.
- Van der Velden A.W., Thomas A.A.M.* The role of 5' untranslated region of an mRNA in translation regulation during development // *Intern. J. Biochem. Cell Biol.* 1999. Vol. 31. P. 87–106.
- Wolberger C.* Combinatorial transcription factors // *Curr. Opin. Genet. Develop.* 1999. Vol. 8. P. 552–559.
- Wolberger C.* Multiprotein–DNA complexes in transcriptional regulation // *Annu. Rev. Biophys. Biomol. Struct.* 1999. Vol. 28. P. 29–56.

К главе 4

- Bell S.P.* Eukaryotic replicators and associated protein complexes // *Curr. Opin. Genet. Develop.* 1995. Vol. 5. P. 162–167.
- Cesareni G., Heimer-Citterich M., Castagnoli L.* Control of ColE1 plasmid replication by antisense RN / . *Trends Genet.* 1991. Vol. 7. P. 230–235.

- Chen C.W.* Complications and implications of linear bacterial chromosomes // Trends Genet. 1996. Vol. 12. P. 192–196.
- Halin J.L., Dijkwel P.A.* On the nature of replication origins in higher eukaryotes // Curr. Opinion Genet. Develop. 1995. Vol. 5. P.153–161.
- Lundblad V., Wright W.E.* Telomeres and telomerase: A simple picture becomes complex // Cell. 1996. Vol. 87. P. 369–375.
- Stillman B.* Smart machines at the DNA replication fork // Cell. 1994. Vol. 78. P. 725–728.

К главе 5

- Ауэрбах Ш.* Проблемы мутагенеза. М.: Мир, 1978. 463 с.
- DNA repair. A special issue. // Trends Biochem. Sci. 1995. Vol. 20. P. 381–440.
- Friedberg E.C., Gerlach V.L.* Novel DNA polymerases offer clues to the molecular basis of mutagenesis // Cell. 1999. Vol. 98. P. 413–416.
- McMurray C.T.* Mechanisms of DNA expansion // Chromosoma. 1995. Vol. 104. P. 2–13.
- Miller J.H.* Spontaneous mutators in bacteria: Insight into pathways of mutagenesis and repair // Annu. Rev. Microbiol. 1996. Vol. 50. P. 625–643.
- Patrushev L.I.* Altruistic DNA. About protective functions of the abundant DNA in the eukaryotic genome and its role in stabilizing genetic information // Biochem. Mol. Biol. Intern. 1997. Vol. 41. P. 851–860.
- Rosenberg S.M.* Mutations for survival // Curr. Opinion Genet. Develop. 1997. Vol. 7. P. 829–834.
- Sancar A.* Excision repair in mammalian cells // J. Biol. Chem. 1995. Vol. 270. P. 15915–15918.
- Sniegowski P.D.* The origin of adaptive mutants: Random or nonrandom? // J. Mol. Evol. 1995. Vol. 40. P. 94–101.

К главе 6

Хесин Р.Б. Непостоянство генома. М.: Наука, 1984. 472 с.

Chun J., Schatz D.G. Developmental neurobiology: Alternative ends for a familiar story? // *Curr. Biol.* 1999. Vol. 9. R251–R253.

К главе 7

Патрушев Л.И. Биосинтез белка в искусственных генетических системах // Проблема белка. М.: Наука, 1995. Т.1 Химическое строение белка. С. 354–478.

Рыбчин В.Н. Основы генетической инженерии. СПб.: СПбГТУ, 1999. 521 с.

Шабарова З.А., Богданов А.А., Золотухин А.С. Химические основы генетической инженерии. М.: Изд-во МГУ, 1994. 219 с.

Bohl D., Heard J.-M. Transcriptional modulation of foreign gene expression in engineering somatic tissues // *Cell. Biol. Toxicol.* 1998. Vol. 14. P. 83–94.

Geisse S., Gram H., Kleuser B. et al. Eukaryotic expression systems: A comparison // *Protein Expression and Purification.* 1996. Vol. 8. P. 271–282.

К главе 8

Chang T.K., Jackson D.Y., Burnier J.P. et al. Subtiligase: a tool for semisynthesis of proteins // *Proc. Nat. Acad. Sci. US.* 1994. Vol. 91. P. 12544–12548.

Houghten R.A. Peptide libraries: Criteria and trends // *Trends Genet.* 1993. Vol. 9. P. 235–239.

Jackson D.Y., Burnier J., Quan C. et al. A designed peptide ligase for total synthesis of ribonuclease A with unnatural catalytic residues // *Science.* 1994. Vol. 226. P. 243–247.

O'Brien P.J., Hershlag D. Catalytic promiscuity and the evolution of new enzymatic activities // *Chem. Biol.* 1999. Vol. 6. P. R91–R105.

Scott J.K. Discovering peptide ligands using epitope libraries // Trends Biochem. Sci. 1992. Vol. 17. P. 241–245.

Wakasugi K., Schimmel P. Two distinct cytokines released from a human aminoacyl-tRNA synthesis // Science. Vol. 284. P. 147–151.

К главе 9

Crooke S.T. Progress in antisense therapeutics // Med. Res. Rev. 1996. Vol. 16. P. 319–344.

Dolnick B.J. Naturally occurring antisense RNA // Pharmacol. Ther. 1997. Vol. 75. P. 179–184.

Famulok M. Oligonucleotide aptamers that recognize small molecules // Curr. Opinion Struct. Biol. 1999. Vol. 9. P. 324–329.

Hager A.J., Pollard J.D., Szostak J.W. Ribozymes: aiming at RNA replication and protein synthesis // Chem. Biol. 1996. Vol. 3. P. 717–725.

Hnatowich D.J. Antisense and nuclear medicine // J. Nucl. Med. 1999. Vol. 40. P. 693–703.

Jones J.T., Lee S.-W., Sullinger B.A. Tagging ribozyme reaction sites to follow *trans*-splicing in mammalian cells // Nature Medicine. 1996. Vol. 2. P. 643–648.

Li Y., Breaker R.R. Phosphorylating DNA with DNA // Proc. Natl. Acad. Sci. (USA). 1999. Vol. 96. P. 2746–2751.

Li Y., Breaker R.R. Deoxiribozymes: new players in the ancient game of biocatalysis // Curr. Opinion Struct. Biol. 1999. Vol. 9. P. 315–323.

Marr J.J. Ribozymes as therapeutic agents // Drug Discov. Today. 1996. Vol. 1. P. 94–102.

Sokol D.L., Murrey J.D. Antisense and ribozyme constructs in transgenic animals // Transgenic Res. 1996. Vol. 5. P. 363–371.

Wilson D.S., Szostak J.W. In vitro selection of functional nucleic acids // Annu. Rev. Biochem. 1999. Vol. 68. 611–647.

К главе 10

- Свердлов Е.Д.* Очерки современной молекулярной генетики по курсу лекций для студентов биологического факультета МГУ. Очерк 5. Трансгеноз и новая молекулярная генетика // Молекуляр. генетика, микробиология и вирусология. 1996. № 4. С. 3–32.
- Свердлов Е.Д.* Очерки современной молекулярной генетики по курсу лекций для студентов биологического факультета МГУ. Очерк 6. Генотерапия и медицина XXI века // Там же. 1997. № 2. С. 3–28.
- Altaner C.* Gene therapy for cancer (present status) // *Neoplasma*. 1995. Vol. 42. P. 209–213.
- Calos M.P.* The potential of extrachromosomal replicating vectors for gene therapy // *Trends Genet.* 1996. Vol. 12. P. 463–466.
- Cristal R.G.* Transfer of genes to humans: Early lessons and obstacles to success // *Science*. 1995. Vol. 270. P. 404–410.
- Harris J.D., Lemoin N.R.* Strategies for targeted gene therapy // *Trends Genet.* 1996. Vol. 12. P. 400–405.
- Mastrangelo M.J., Berd D., Nathan F.E.* et al. Gene therapy for human cancer: An assay for clinicians // *Semin. Oncol.* 1996. Vol. 23. P. 4–21.
- Morishita R., Higaki J., Tomita N., Ogiwara T.* Application of transcription factor "decoy" strategy as means of gene therapy and study of gene expression in cardiovascular disease // *Circ. Res.* 1999. Vol. 82. P. 1022–1028.
- Shastri B.S.* Genetic knockouts in mice: An update // *Experientia*. 1995. Vol. 51. P. 1028–1039.
- Spencer D.M.* Creating conditional mutations in mammals // *Trends Genet.* 1996. Vol. 12. P. 181–187.
- Weiss B., Davidkova G., Zhou L.-W.* Antisense RNA gene therapy for studying and modulating biological processes // *Cell Mol. Life Sci.* 1999. Vol. 55. P. 334–358.
- Zanjani E.D., Anderson W.F.* Prospects for in utero human gene therapy // *Science*.

1999. Vol. 285. P. 2084–2088.

К главе 11

Шагинян И.А., Гинцбург А.Л. ПЦР-генетическое типирование патогенных микроорганизмов // Генетика. 1995. Т. 31. С. 600–610.

Graber J.H., O'Donnell M.J., Smith C.L., Cantor C.R. Advances in DNA diagnostics // Curr. Opinion Biotechnol. 1998. Vol. 9. 14–18.

Jeffreys A.J., MacLeod A., Tamaki K. et al. Minisatellite repeat coding as digital approach to DNA typing // Nature. 1991. Vol. 354. P. 204–209.

Nuwaysir E.F., Bittner M., Trent J., Barret C., Afshary C.A. Microarrays and toxicology: The advent of toxicogenomics // Mol. Carcinogenesis. 1999. Vol. 24. P. 153–159.

Parsons B.L., Heflich R.H. Genotypic selection methods for the direct analysis of point mutations // Mutation Res. 1997. Vol. 387. P. 97–121.

Pena S.D.J., Chakraborty R. Paternity testing in the DNA era // Trends Genet. 1994. Vol. 10. P. 204–209.

Roth M., Yarmush M.L. Nucleic acid biotechnology // Annu. Rev. Biomed. Eng. 1999. Vol. 1. P. 265–297.

Shikata H., Utsumi N., Kuivaniemi H. et al. DNA-based diagnostics in the study of heritable and acquired disorders // J. Lab. Clin. Med. 1995. Vol. 125. P. 421–432.

Sidransky D. Nucleic acid-based for the detection of cancer // Science. 1997. Vol. 278. P. 1054–1058.

К главе 12

Benner S.A., Trabesinger N., Schreiber D. Past-genomic science: Converting primary structure into physiological function // Adv. Enzyme Regul. 1998. Vol. 38. P. 155–180.

Billings P.R., Smith C.L., Cantor C.L. New techniques for physical mapping of the

human genome // FASEB J. 1991. Vol. 5. P. 28–34.

Burge C.B., Karlin S. Finding the genes in genomic DNA // Curr. Opin. Struct. Biol. 1998. Vol. 8. P. 346–354.

Clark M.S. Comparative genomics: The key to understanding the Human Genome Project // BioEssays. 1999. Vol. 21. P. 21–30.

McKusick V. Current trends in mapping human genes // FASEB J. 1991. Vol. 5. P. 12–20.

Rose E.A. Applications of the polymerase chain reaction to genome analysis // Ibid. 1991. Vol. 5. P. 46–54.

Научное издание

Патрушев Лев Иванович

Экспрессия генов

Утверждено к печати Ученым советом
Института биоорганической химии
им. М.М. Шемякина и Ю.А. Овчинникова
Российской академии наук

Заведующая редакцией А.М. Гидалевич
Редактор Т.И. Белова

ЛР № 020297 от 23.06.1997